THE PROCEEDINGS OF THE PHYSICAL SOCIETY

Section B

Vol. 64, Part 8 1 August 1951 No. 380 B

CONTENTS

PAGE

Dr. M. Blackman. Diffraction from a Bent Crystal 625

Dr. M. Blackman. Diffraction from a Curved Linear Lattice 631

Dr. J. M. Cowley, Dr. A. L. G. Rees and Mr. J. A. Spink. The Morphology of Zinc Oxide Smoke Particles 638

Dr. J. Corner and Mr. E. D. Pendlebury. The Coagulation and Deposition of a Stirred Aerosol 645

Dr. R. Q. Twiss. On Oscillations in Electron Streams 654

Dr. R. E. B. Makinson, Dr. P. C. Thonemann, Mr. R. B. King and Mr. J. V. Ramsay. Dielectric Constant and Electron Density in a Gas Discharge . . 665

Dr. K. W. Plessner. Conductivity, Hall Effect and Thermo-electric Power of Selenium Single Crystals 671

Dr. K. W. Plessner. Conductivity and Hall Effect of Micro-crystalline Selenium containing Iodine Impurities 681

Dr. C. A. Hogarth. Some Conduction Properties of the Oxides of Cadmium and Nickel 691

Dr. W. Ehrenberg and Dr. J. Hirsch. The Conductivity–Temperature Characteristics of Lead Sulphide: The Influence of Oxygen and of the Rate of Heating. 700

Dr. R. Bechmann and Mr. P. L. Parsons. Piezoelectric Behaviour of Partially Plated Square Plates Vibrating in Contour Modes 706

Dr. R. S. Longhurst. The Astigmatic Correction of Telescopes 713

Letters to the Editor:

  Dr. J. A. Clegg and Mr. R. L. Closs. Plasma Oscillations in Meteor Trails . 718

  Dr. J. M. Cassels, Mr. J. M. Dickson and Dr. J. Howlett. Multiple Traversal of Cyclotron Targets: an Extension to the Theory 719

Reviews of Books 720

Contents for Section A 726

Abstracts for Section A 727

Price to non-members 10*s*. net, by post 6*d*. extra. Annual subscription: £5 5*s*.
Composite subscription for both Sections A and B: £9 9*s*.

Published by
THE PHYSICAL SOCIETY
1 Lowther Gardens, Prince Consort Road, London S.W.7

PROCEEDINGS OF THE PHYSICAL SOCIETY

The *Proceedings* is now published monthly in two Sections.

ADVISORY BOARD

Chairman : The President of the Physical Society (L. F. Bates, D.Sc., Ph.D., F.R.S.)

E. N. da C. Andrade, Ph.D., D.Sc., F.R.S.
Sir Edward Appleton, G.B.E., K.C.B., D.Sc., F.R.S.
P. M. S. Blackett, M.A., F.R.S.
Sir Lawrence Bragg, O.B.E., M.A., Sc.D., D.Sc., F.R.S.
Sir James Chadwick, D.Sc., Ph.D., F.R.S.
S. Chapman, M.A., D.Sc., F.R.S.
Lord Cherwell of Oxford, M.A., Ph.D., F.R.S.
Sir John Cockcroft, C.B.E., M.A., Ph.D., F.R.S.
Sir Charles Darwin, K.B.E., M.C., M.A., Sc.D., F.R.S.
N. Feather, Ph.D., F.R.S.
G. I. Finch, M.B.E., D.Sc., F.R.S.
D. R. Hartree, M.A., Ph.D., F.R.S.
N. F. Mott, M.A., F.R.S.
M. L. Oliphant, Ph.D., D.Sc., F.R.S.
F. E. Simon, C.B.E., M.A., D.Phil., F.R.S.
T. Smith, M.A., F.R.S.
Sir George Thomson, M.A., D.Sc., F.R.S.

Papers for publication in the *Proceedings* should be addressed to the Hon. Papers Secretary, Dr. H. H. Hopkins, at the Office of the Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7. Telephone : KENsington 0048, 0049.

Detailed Instructions to Authors were included in the February 1948 issue of the *Proceedings*; separate copies can be obtained from the Secretary-Editor.

BULLETIN ANALYTIQUE

Publication of the Centre National de la Recherche Scientifique, France

The *Bulletin Analytique* is an abstracting journal which appears in three parts, Part 1 covering scientific and technical papers in the mathematical, chemical and physical sciences and their applications, Part 2 the biological sciences and Part 3 philosophy.

The *Bulletin*, which started on a modest scale in 1940 with an average of 10,000 abstracts per part, now averages 35 to 45,000 abstracts per part. The abstracts summarize briefly papers in scientific and technical periodicals received in Paris from all over the world and cover the majority of the more important journals in the world scientific press. The scope of the *Bulletin* is constantly being enlarged to include a wider selection of periodicals.

The *Bulletin* thus provides a valuable reference book both for the laboratory and for the individual research worker who wishes to keep in touch with advances in subjects bordering on his own.

A specially interesting feature of the *Bulletin* is the microfilm service. A microfilm is made of each article as it is abstracted and negative microfilm copies or prints from microfilm can be purchased from the editors.

The subscription rates per annum for Great Britain are 4,000 frs. (£4) each for Parts 1 and 2, and 2,000 frs. (£2) for Part 3. Subscriptions can also be taken out to individual sections of the *Bulletin* as follows :

	frs.	
Pure and Applied Mathematics—Mathematics—Mechanics	550	14/6
Astronomy—Astrophysics—Geophysics	700	18/–
General Physics—Thermodynamics—Heat—Optics—Electricity and Magnetism	900	22/6
Atomic Physics—Structure of Matter	325	8/6
General Chemistry—Physical Chemistry	325	8/6
Inorganic Chemistry—Organic Chemistry—Applied Chemistry—Metallurgy	1,800	45/–
Engineering Sciences	1,200	30/–
Mineralogy—Petrography—Geology—Palaeontology ..	550	14/6
Biochemistry—Biophysics—Pharmacology	900	22/6
Microbiology—Virus and Phages	600	15/6
Animal Biology—Genetics—Plant Biology	1,800	45/–
Aggriculture—Nutrition and the Food Industries ..	550	14/6

Subscriptions can be paid directly to the editors : Centre National de la Recherche Scientifique, 18, rue Pierre-Curie, Paris 5ème (Compte-chèque-postal 2,500–42, Paris), or through Messrs. H. K. Lewis & Co. Ltd., 136, Gower Street, London W.C.1.

PROCEEDINGS OF THE PHYSICAL SOCIETY
in
MICROFILM

The Physical Society has agreed with University Microfilms, Ann Arbor, Michigan, for the reproduction of the *Proceedings of the Physical Society* in Microfilm form.

This service is available only to subscribers to the paper edition of the Journal, and is intended to be of assistance to libraries both in saving accessible space and in improving borrowing facilities.

The microfilm is produced as a 'positive', i.e. black printing on white background, and is supplied on metal reels suitably labelled, distribution being made at the end of the year.

Inquiries to be addressed to
THE UNIVERSITY MICROFILMS
313 N. First Street, Ann Arbor, Michigan, U.S.A.

PAST ISSUES
OF THE
PROCEEDINGS OF THE PHYSICAL SOCIETY
AND THE
TRANSACTIONS OF THE OPTICAL SOCIETY

Your attention is drawn to the fact that **Messrs. Wm. Dawson & Sons Ltd.,** 102 Wigmore Street, London W.C. 1, are now acting as agents for all issues of the *Proceedings of the Physical Society* up to and including 1948, and the *Transactions of the Optical Society*, Volumes 1–33.

Orders for these publications should be addressed to Messrs. Wm. Dawson direct.

The current volume and the two previous years of the *Proceedings* and all special publications are obtainable from the **Offices of the Society** in the normal way.

HANDBOOK
OF THE
PHYSICAL SOCIETY'S 35th EXHIBITION
OF
SCIENTIFIC INSTRUMENTS AND APPARATUS
1951

lxxii + 244 pp.; 121 illustrations
5*s*.; *by post* 6*s*.

Orders, with remittances, to
THE PHYSICAL SOCIETY
1 Lowther Gardens, Prince Consort Road,
London S.W.7

PROCEEDINGS OF THE PHYSICAL SOCIETY

ADVERTISEMENT RATES

The *Proceedings* are divided into two parts, A and B. The charge for insertion is £18 for a full page in either Section A or Section B, £30 for a full page for insertion of the same advertisement in both Sections. The corresponding charges for part pages are :

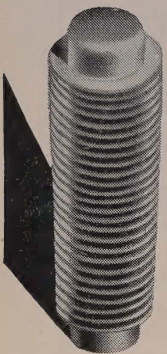
½ page	£9 5 0	£15 10 0
¼ page	£4 15 0	£8 0 0
⅛ page	£2 10 0	£4 5 0

Discount is 20% for a series of six similar insertions and 10% for a series of three.

The printed area of the page is 8½″ × 5½″, and the screen number is 120.

Copy should be received at the Offices of the Physical Society six weeks before the date of publication of the *Proceedings*.

SEAMLESS ONE-PIECE METAL BELLOWS



Combining the properties of:

1 A compression spring capable of repeated flexing

2 A container which can be hermetically sealed

3 A packless gland

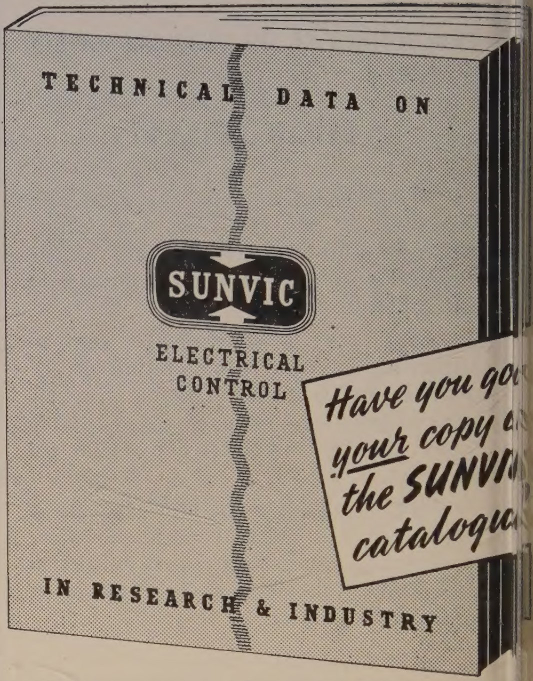
Hydraulically formed by a process unique in this country

for Automatic coolant regulation. Movement for pressure change. Packless gland to seal spindle in high vacua. Reservoir to accept liquid expansion. Dashpot or delay device. Barometric measurement or control. Pressurised couplings where vibration or movement is present. Dust seal to prevent ingress of dirt. Pressure reducing valves. Hydraulic transmission. Distance thermostatic control. Low torque flexible coupling. Pressure sealed rocking movement. Pressurised rotating shaft seals. Aircraft pressurised cabin control. Refrigeration expansion valves. Thermostatic Stream Traps. Pressure amplifiers. Differential pressure measurements. Thermostatic operation of louvre or damper. *Write for List No.* V. 800–1.

by DRAYTON

B10

Drayton Regulator & Instrument Co. Ltd., West Drayton, Mdx. · W. Drayton 2611



Precision temperature controls include: Thermostats—Nc Energy Regulators—Hotwire Vacuum Switches—Time D —Electronic and control apparatus.

Write today to SUNVIC CONTROLS L

Member of the A.E.I. Group of Companies

SUNVIC HOUSE, 10 ESSEX ST., LONDON, W.

THE PHYSICAL SOCIETY

VOLUME XIV of the REPORTS ON PROGRESS IN PHYSICS

A comprehensive annual review by specialist authors. The contents are as follows:

W. C. Price. Recent Advances in Ultra-Violet Absorption Spectroscopy.
W. E. Lamb, Jr. Anomalous Fine Structure of Hydrogen and Singly Ionized Helium.
H. Kuhn. New Techniques in Optical Interferometry.
E. Wolf. Diffraction Theory of Aberrations.
A. B. Meinel. The Spectrum of the Airglow and the Aurora.
B. J. Mason and F. H. Ludlam. The Microphysics of Clouds.
M. Dfutsch. Angular Correlations in Nuclear Reactions.
G. D. Rochester and W. V. G. Rosser. Nuclear Interactions of Cosmic Rays.
E. W. Foster. Nuclear Effects in Atomic Spectra.
N. C. Gerson. A Critical Survey of Ionospheric Temperatures.
W. V. Mayneord. Some Applications of Nuclear Physics in Medicine.

The price is 50*s*. 0*d*. Members: One copy at 27*s*. 6*d*.
Postage and packing 1*s*.

Further information can be obtained from

THE PHYSICAL SOCIETY

1 Lowther Gardens, Prince Consort Road, London S.W.7

Diffraction Gratings

Hilger & Watts Ltd. are about to produce plane replica gratings of various kinds based on the methods of Sir Thomas Merton (see *Proc. Roy. Soc.* A, **201**, 187, "On the reproduction and ruling of diffraction gratings") and using the replica technique as developed by the National Physical Laboratory (see *Proc. Roy. Soc.* A, **207**, 278, "On the production of diffraction gratings—Pt. 1: The copying of plane gratings", by G. H. Dew & L. A. Sayce). They will also be able to supply either replica or original concave gratings.

It will help them to assess the priority of production if those who are likely to require gratings, either soon or in the future, will write, indicating the line spacing, area of ruling, concentration in a particular direction, and any other special features. Such information will not be taken to imply any commitment to purchase such gratings.

The following original gratings are **available now** from stock subject to prior sale. Spacing, 14,400 lines per inch (567 lines per millimetre). Ruled area up to 8 × 5 centimetres. Plane, or concave with radii 1, 2, or 3 metres.

HILGER & WATTS LTD: HILGER DIVISION
98 St. Pancras Way, Camden Rd., London, N.W.1
Telephone: GULliver 5571

Member of Scientific Instrument Manufacturers' Association & SCIEX

HILGER

PHYSICAL SOCIETY SPECIALIST GROUPS

OPTICAL GROUP

The Physical Society Optical Group exists to foster interest in and development of all branches of optical science. To this end, among other activities, it holds meetings about five times a year to discuss subjects covering all aspects of the theory and practice of optics, according to the papers offered.

COLOUR GROUP

The Physical Society Colour Group exists to provide an opportunity for the very varied types of worker engaged on colour problems to meet and to discuss the scientific and technical aspects of their work. Five or six meetings for lectures and discussions are normally held each year, and reprints of papers are circulated to members when available. A certain amount of committee work is undertaken, and reports on Defective Colour Vision (1946) and on Colour Terminology (1948) have already been published.

LOW TEMPERATURE GROUP

The Low Temperature Group was formed to provide an opportunity for the various groups of people concerned with low temperatures—physicists, chemists, engineers, etc.—to meet and become familiar with each other's problems. The group seeks to encourage investigations in the low temperature field and to assist in the correlation and publication of data.

ACOUSTICS GROUP

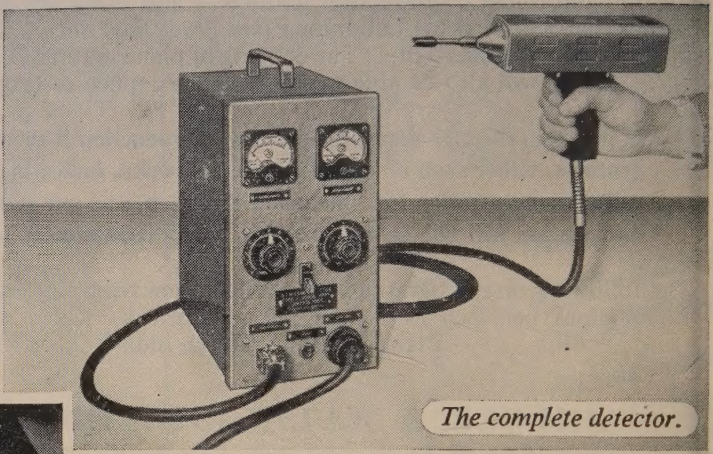
The Acoustics Group was formed to meet the long-felt need for a focus of acoustical studies in Great Britain. The scope includes the physiological, architectural, psychological and musical aspects of acoustics as well as the fundamental physical studies on intensity transmission and absorption of sound. The Group achieves its object by holding discussion meetings, by the circulation of reprints and by arranging symposia on selected acoustical topics.

Further information may be obtained from the Offices of the Society:
1 Lowther Gardens, Prince Consort Road, London S.W.7.

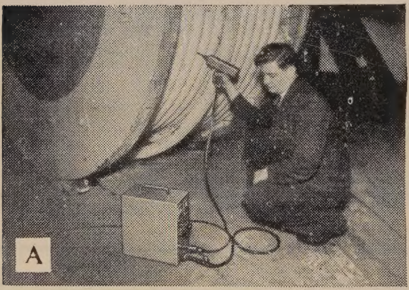
B T-H

Electronic
LEAK DETECTOR

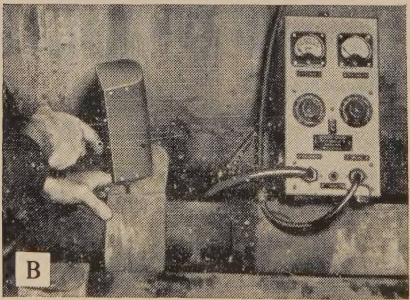
Convenient
Accurate
Sensitive



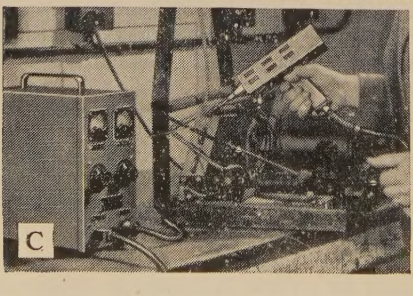
The complete detector.



A



B



C

Leaks in containers or pipes imperceptible by other means can be easily and accurately located by the BTH Type H Electronic Leak Detector. This instrument, developed in the BTH Research Laboratories to detect the vapours of halogen compounds in air, is sensitive enough to react to Arcton ($C\,CL_2\,F_2$) escaping from a container at the rate of 1/50 oz. per year. The detector is easily portable—eminently suited to use in the laboratory or on the assembly line, or for service testing in the field. Applications include testing of refrigeration and air-conditioning plant, which already contains a halogen-bearing compound, and leak-testing of tanks, pipes, joints, welds, pneumatic systems, etc. by introducing a suitable " tracer " gas into the system under test. The detector emits, by means of a small built-in sounder, a series of clicks ; a change of frequency denotes the presence of a leak. For work in noisy surroundings headphones are available.

Illustrations show use of the detector in examining for leaks (A) *Gas-filled cables* (B) *Pipe seams* (C) *Refrigerator unit.*

THE
BRITISH THOMSON-HOUSTON
COMPANY LIMITED, RUGBY, ENGLAND

Member of the AEI group of companies

A4293

THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

Section B

VOL. 64, PART 8 1 August 1951 No. 380 B

Diffraction from a Bent Crystal

BY M. BLACKMAN
Physics Department, Imperial College, London

MS. received 11*th January* 1951

ABSTRACT. The diffraction of plane waves by a bent crystal is considered. It is shown that the peak of the diffracted ' intensity ' can be represented in reciprocal space by an arc, the limits of the arc being the total angle of bend of the crystal. The ' intensity ' function in reciprocal space has been studied numerically in a particular case, and it is shown that both ' intensity ' and shape change with order.

§ 1

THE theory of the diffraction of a beam of parallel rays by a bent crystal has so far received very little attention. The theory is of interest in connection, among other things, with the interpretation of electron diffraction transmission photographs of certain specimens of mica and molybdenite (Darbyshire and Cooper 1935, Finch and Wilman 1936); these show features which are very similar to, if not identical with, the photographs obtained by the rotation of a normal specimen. It has been commonly assumed that a bent crystal will give such 'rotation' patterns, but no theoretical investigation has proved this, nor is there any means of finding the detailed form of the intensity of the diffracted wave. The only serious investigation which the author has been able to find is that due to Fock and Kolpinsky (1940), who investigated the diffraction from a two-dimensional lattice bent into the form of a circular cylinder; this really amounts to the calculation of the diffraction from points spaced uniformly on the circumference of a circle. The calculation does not lead directly to the diffraction effects from a three-dimensional lattice, and the authors eventually fall back on the physical arguments, quoted above, which have been used before.

In the present calculation a simple case is considered, the bent lattice consisting of a set of bent planes placed at a fixed distance apart.

§ 2.

The total amplitude of the waves scattered from a set of scattering points is proportional to the quantity

$$G = \sum_k \exp\left[2\pi i(\mathbf{r}_k \cdot (\mathbf{s}-\mathbf{s}_0)/\lambda)\right]. \qquad \ldots\ldots(1)$$

Here $\mathbf{r}_k$ is a vector from an arbitrary origin to one of the scattering centres; $\mathbf{s}_0$ is a unit vector specifying the incident place wave (i.e. parallel to the normal to the wave front), and $\mathbf{s}$ is the corresponding vector for the diffracted wave. The above relation can be applied to the bent lattice described above. The form of the bend is shown in Figure 1; the curved lines have a common centre of curvature, the distance between the lines being taken to be fixed and equal to a. This is also taken to be the distance between points on a particular arc (the 'reference arc'), the distance a being measured along the arc.

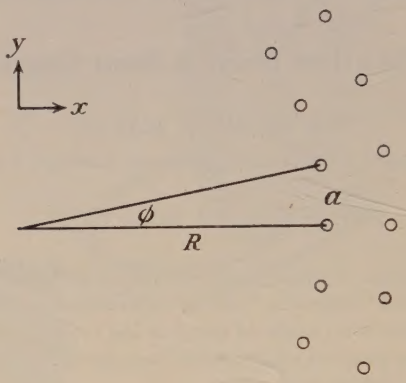


Figure 1.

The crystal is formed by the addition of similar layers at a fixed distance apart. In this case (1) can be written as

$$G = L \sum_{ql} \exp\,[2\pi i\{\mathbf{r}_{ql}\,.\,(\mathbf{s}-\mathbf{s}_0)\}/\lambda], \qquad \ldots\ldots(2)$$

where the double summation refers to the plane shown in Figure 1. The term L contains the summation over the third direction, and, except for an irrelevant phase factor, has the familiar form $L=\sin[N_3A_3/2]/\sin\,[A_3/2]$, where $A_3=2\pi\{\mathbf{a}_3\,.(\mathbf{s}-\mathbf{s}_0)\}/\lambda$ and $|\mathbf{a}_3|$ is the spacing between planes; N_3 is the total number of planes. The vector $\mathbf{r}_{ql}$ is split into two parts $\mathbf{r}_{ql}=\mathbf{R}_l+\mathbf{r}_l{}^q$, where $\mathbf{R}_l$ is the vector from the origin (the centre of curvature in a particular plane) to the point l of the reference arc; $\mathbf{r}_l{}^q$ is the vector from this point to any point on the radial line from the origin through l; these are numbered by q at fixed l. Using the above description of the bent crystal we can write $\mathbf{r}_l{}^q=qa\mathbf{i}_l$ where $\mathbf{i}_l$ is a unit vector. The components of this vector are $\mathbf{i}_{lx}=\cos l\phi$, $\mathbf{i}_{ly}=\sin l\phi$, where the x axis is the line from the origin to the zero point ($l=0$) of the reference arc; the angle ϕ is a/R where $R=|\mathbf{R}_l|$. Following the notation used by Fock and Kolpinsky (1940) R is expressed as $R=pa/2\pi$ and hence $\phi=2\pi/p$. With the above rearrangements we can write the double sum in (2) as

$$S=\sum_l \exp\,[2\pi i\{\mathbf{R}_l\,.\,(\mathbf{s}-\mathbf{s}_0)\}/\lambda]\,.\sum_q \exp\,[2\pi iqa\{\mathbf{i}_l\,.\,(\mathbf{s}-\mathbf{s}_0)\}/\lambda]. \qquad \ldots\ldots(3)$$

Writing $(\mathbf{s}-\mathbf{s}_0)/\lambda=\mathbf{H}$ where $\mathbf{H}$ is a vector in reciprocal space, this becomes

$$S=\sum_l \exp\,[2\pi i(\mathbf{R}_l\,.\,\mathbf{H})]\,\sin\,(\tfrac{1}{2}N_1\psi_l)/\sin\,(\tfrac{1}{2}\psi_l) \qquad \ldots\ldots(4)$$

where $\psi_l=2\pi(aH_x\cos 2\pi l/p+aH_y\sin 2\pi l/p)$ and N_1 is the total number of curved lines in the plane (i.e. the total number of q values). This form is made possible because the summation over q in (3) is taken at a fixed value of l. The summation over q which gives the form (4) is taken from $-(N_1-1)/2$ to $+(N_1-1)/2$. The reference arc is characterized by $q=0$.

We can also express the first part of (4) containing $(\mathbf{R}_l\,.\,\mathbf{H})$ in terms of ψ_l:

$$2\pi(\mathbf{R}_l\,.\,\mathbf{H})=p(aH_x\cos l\phi+aH_y\sin l\phi)=p\psi_l/2\pi.$$

Hence $$S=\sum_l \exp(ip\psi_l/2\pi)\sin(\tfrac{1}{2}N_1\psi_l)/\sin(\tfrac{1}{2}\psi_l), \qquad \ldots\ldots(5)$$

the summation being taken from $-l'$ to $+l'$. The calculation of the diffracted intensity then depends on that of $|S|^2$ since $|G|^2=|\mathbf{L}|^2\,.\,|\mathbf{S}|^2$.

It will be seen that the ratio of the sine terms in (5) has a peak value when $\psi_{l_0}=2m\pi$ or

$$aH_x\cos 2\pi l_0/p+aH_y\sin 2\pi l_0/p=m \qquad \ldots\ldots(6)$$

where m is an integer. This suggests that ψ_l should be developed round the point $l=l_0$. We can write

$$\psi_l/2\pi=\cos 2\pi(l-l_0)/p\,[aH_x\cos 2\pi l_0/p+aH_y\sin 2\pi\,l_0/p]$$
$$+\sin 2\pi(l-l_0)/p\,[-aH_x\sin 2\pi l_0/p+aH_y\cos 2\pi\,l_0/p].$$

In all cases in which we are likely to be interested, p will be large compared with l'. We expand the sine and cosine terms and can write

$$p\psi_l/2\pi=[aH_x\cos 2\pi l_0/p+aH_y\sin 2\pi l_0/p][p-2\pi^2(l-l_0)^2/p+\ \ldots\ldots]$$
$$+[-aH_x\sin 2\pi\,l_0/p+aH_y\cos 2\pi\,l_0/p][2\pi(l-l_0)-4\pi^3(l-l_0)^3/3p^2+\ldots\ldots]. \qquad \ldots(7)$$

It will be seen that if

$$-aH_x\sin 2\pi l_0/p+aH_y\cos 2\pi l_0/p=n, \qquad \ldots\ldots(8)$$

where n is integral, $\exp(ip\psi_l/2\pi)$ depends on terms in $(l-l_0)$ higher than the first power if we take (6) into account. For, since l is integral, the terms in l will give a factor of unity; if l_0 is not integral this introduces a phase factor into S as does the first term in (7) which contains p and not $(l-l_0)$. Since we are interested in $|\mathbf{S}|^2$, these phase factors are unimportant.

The two conditions (6) and (8) ensure a point of 'stationary phase' at $l=l_0$, and it is to be expected that these conditions will determine those values of H_x, H_y for which $|\mathbf{S}|^2$ will be large.

These conditions of stationary phase, which replace the usual Laue conditions, can be put into a clearer form if polar coordinates are used in reciprocal space. Putting $aH_x=aH_p\sin\alpha=\rho\sin\alpha$, $aH_y=aH_p\cos\alpha=\rho\cos\alpha$, where H_p is the component of $\mathbf{H}$ in the xy plane, and ρ is dimensionless, the conditions (6), (8) become

$$(a)\ \rho\cos(2\pi l_0/p+\alpha)=n, \qquad (b)\ \rho\sin(2\pi l_0/p+\alpha)=m \qquad \ldots\ldots(9)$$

or $$(a)\ \rho^2=m^2+n^2, \qquad (b)\ \alpha=\tan^{-1}(m/n)-2\pi l_0/p. \qquad \ldots\ldots(10)$$

The first condition (10a) defines a circle in reciprocal space passing through the 'Laue point' $(H_x=n/a,\ H_y=m/a)$. The second condition limits this circle to an arc, since l_0 can vary from $-l'$ to $+l'$. This arc is symmetrically situated with respect to the Laue point, and the total angle subtended by the arc at the origin is $2\pi(2l'+1)/p$, which is the total angle of the bend.

This is a theoretical justification, as far as it goes, for the commonly held belief that the effect of bending a crystal is to draw out a Laue point into an arc in the reciprocal lattice. For the limiting case $p=\infty$, the above conditions reduce to those for a Laue point.

There is of course a third condition, which derives from the term L in (2), and which is the normal Laue condition $(\mathbf{a}_3 \,.\, \mathbf{H}) = k$. We shall be discussing the case of $k = 0$ only, but the general form of $|G|^2$ in reciprocal space can be immediately deduced from the form of L given above.

§ 3

The next stage consists in showing how far the above 'conditions of stationary phase' represent a region of high intensity in reciprocal space. To do this we shall have to calculate $|\mathbf{S}|^2$ as a function of H_x, H_y in this region. This is done by numerical integration, the summation being split into its real and imaginary parts. The calculations performed so far have been restricted to points in reciprocal space lying in a radial line passing through the Laue points. These will not differ for other radial lines until the edge of the permitted arc is reached. In all cases it has been ascertained that the conversion from a sum to an integral is justified and the contribution of the higher terms in $(l - l_0)$ have been allowed for where necessary.

When the values ρ, α differ from those which satisfy conditions $(9\,a, b)$ we have a certain amount of freedom in handling the summation. In most cases it seems best to choose l_0 so as to keep condition $(9\,b)$, i.e. $\rho \sin(2\pi l_0/p + \alpha) = m$, then $\rho \cos(2\pi l_0/p + \alpha) = (\rho^2 - m^2)^{1/2}$. Since the calculation is carried through at fixed α and varying ϵ, where ϵ is given by $\rho = (m^2 + n^2)^{1/2}(1 + \epsilon)$, the above can be written $\rho \cos(2\pi l_0/p + \alpha) = n[1 + (n^2 + m^2)\epsilon/n^2]$ where $\epsilon \ll 1$, as will be the case for the numerical calculations described below. These values are inserted in the expression for S, and the calculations carried out as for $\epsilon = 0$.

In certain special cases other formulations must be used. For instance if $n = 0$ the condition $(9\,a)$ is the obvious one to start with: $\rho \cos(2\pi l_0/p + \alpha) = 0$, which fixes l_0, and $\rho \sin(2\pi l_0/p + \alpha) = m(1 + \epsilon)$.

The above methods can be used for all values of α inside those specified by (10). The effect of the limiting values of l must however be watched. Outside the region given by (10) other methods would have to be used.

In the detailed investigation l' is taken as 10^6, N_1 as 2,000. The value of p is chosen to be $3{\cdot}6 \times 10^7$, giving a total angle of bend of 20°. These numbers are chosen so as to correspond roughly with those pertaining to thin films of mica used in electron diffraction. These films show diffraction effects which have been interpreted as due to bending. The value of N_3 need not be specified, as it merely fixes the width of the intensity region in a direction in reciprocal space perpendicular to the plane containing the arc.

Three orders are considered, the (m, n) values being $(0, 1)$ $(1, 1)$ $(1, 4)$. The value of $|\mathbf{S}|^2$ for these cases are shown in Figure 2 as ordinates, the abscissae are distances in reciprocal space in units of a^{-1} measured from the Laue point. Though $|\mathbf{S}|^2$ is not strictly speaking symmetrical about the Laue point the difference is too small, in the above cases, to be indicated in the diagrams and hence only one half of the curve is shown.

The features to be noted are the width and structure of the curves, which vary from order to order. This is indeed the main difference between the bent lattice and an unstrained lattice rotated through an angle. In the latter case the structure of $|\mathbf{S}|^2$ would remain unaltered with order.

The shape of the curves will vary considerably with the numbers chosen for N_1 and l'. If we fix l', that is keep the angle of bend constant (p is kept the same), we can obtain some of the main features from elementary arguments. For small N_1 values we would expect this number to control the width of the region of high $|\mathbf{S}|^2$ values in the reciprocal lattice, this being roughly $(N_1 a)^{-1}$, e.g. in the case considered a width of $5 \times 10^{-4} a^{-1}$ is approximately correct for the first order. In addition, for very small N_1, $|\mathbf{S}|^2$ must approach the limit for a bent chain. Here the peaks are very definitely unsymmetrical and have the form of the square of a high order Bessel function (Fock and Kolpinsky 1940). Also the 'atomic' scattering factor, neglected in (1), would have to be inserted. Apart from the value of N_1 there is another factor which influences the width of the curves; this is the change of the spacing of points on a curved line. This, as shown in Figure 1, varies from line to line. The maximum relative change of spacing in the case calculated above is $2 \times 10^3 / 6 \times 10^6 \sim 3 \times 10^{-4}$ and we should expect a width of this order (measured in units of a^{-1}) due to this cause for low orders. If, however, N_1 is increased from 2×10^3 to 2×10^4, the relative change in spacing of the order of 3×10^{-3} will be the main factor in determining the width. There will therefore be a minimum width as N_1 is varied, which will occur when the effects due to distortion and to the finiteness of the lattice are approximately equal.

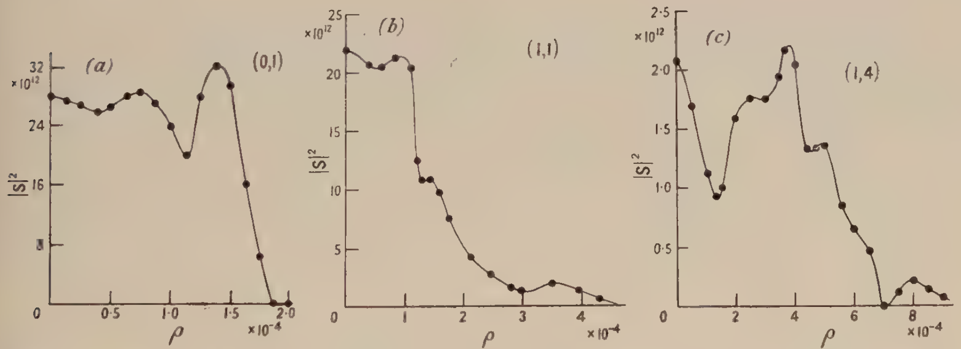


Figure 2. Value of $|\mathbf{S}|^2$ as a function of the radial distance in the reciprocal space, measured from the Laue point (m, n). The (m, n) values are (*a*) (0, 1), (*b*) (1, 1), (*c*) (1, 4). The unit of distance is a^{-1}.

A further general feature of the regions of large $|\mathbf{S}|^2$ values can be seen by referring to the initial expression for S: $S = \Sigma_{ql} \exp[2\pi i(\mathbf{r}_{ql} \cdot \mathbf{H})]$. From this it follows that $S(H_x, H_y) = S^*(-H_x, -H_y)$ and $|\mathbf{S}(H_x, H_y)|^2 = |\mathbf{S}(-H_x, -H_y)|^2$, i.e. the origin is a point of inversion. Also since the lattice is symmetrical on reflection in the x-axis, $|\mathbf{S}(-H_y, H_x)|^2 = |\mathbf{S}(H_y, H_x)|^2$ and it follows that

$$|\mathbf{S}(-H_y, H_x)|^2 = |\mathbf{S}(-H_y, -H_x)|^2.$$

The pattern in the xy plane is symmetrical on reflection in the x and y axes. That for $|\mathbf{G}|^2$ is symmetrical on reflection in the xz, yz planes as well as in the xy plane.

§ 4

The calculations discussed above are restricted in scope in that certain simplifying assumptions were made initially. In particular, since a crystal has elastic properties, the bending will cause a change of spacing between the curved lines (Figure 1). The effect of this is to alter the sum in (3)

$$\Sigma_q \exp\left[2\pi i(\mathbf{r}_l{}^q \cdot \mathbf{H})\right]$$

which gave rise to the term $\sin(\frac{1}{2}N_1\psi_l)/\sin(\frac{1}{2}\psi_l)$. This sum cannot be evaluated in general in such a simple manner, and the calculation is correspondingly more complicated. It is not expected that this will alter the main conclusions of §§2 and 3.

So far only one form of curvature has been introduced into the calculation; the introduction of two finite principal radii of curvature would produce difficulties of the same type as those mentioned in the preceding paragraph. If one ignores these difficulties the main result is to change the arc-shaped region in reciprocal space to a curved lamina bounded by arcs determined from the radii of curvature as in §2. Sections of this lamina by the Ewald sphere would give rise to a streak which would appear as such on a photographic plate suitably placed. In fact the whole of the electron diffraction pattern of such a bent crystal would be a set of such streaks; so far no evidence has been published which shows such phenomena.

§ 5

Whether the theory of the bent lattice can be used to explain the electron diffraction photographs of certain specimens of mica and molybdenite, must be left to a careful investigation of experimental results. The data published so far show that the patterns can be reproduced by a suitable rotation of an undeformed crystal. This does not distinguish between a bent crystal and a crystal originally bent which has split up into blocks aligned more or less in their original positions in the bent crystal. Indeed an exact correspondence with a 'rotation' photograph would prove the latter possibility. An indication in favour of a 'block' theory, for which I am indebted to Dr. W. Cochrane, is to be found in the sharpness of many of the spots in photographs of 'bent' crystals: these appear to be sharper than the spot produced by the main beam. If this is confirmed, it would show that scattering took place from individual 'blocks' smaller in size than the main beam. Another, rather curious, feature which needs further investigation is the change in the crystal structure of mica which Fock and Kolpinsky found necessary to postulate in order to index their patterns from 'bent' mica. The angle of the monoclinic cell has to be changed from 98° 36′ to approximately 110°, which is a somewhat unusual effect of a deformation.

ACKNOWLEDGMENTS

The author wishes to express his thanks to Dr. W. Cochrane for many discussions, and to Sir George Thomson for his interest in this work.

REFERENCES

Darbyshire, J. A., and Cooper, E. R., 1935, *Proc. Roy. Soc.* A, **152,** 104.
Finch, G. I., and Wilman, H., 1936, *Trans. Faraday Soc.*, **32,** 1539.
Fock, V. A., and Kolpinsky, V. A., 1940, *J. Phys. U.S.S.R.*, **3,** 125.

Diffraction from a Curved Linear Lattice

By M. BLACKMAN
Physics Department, Imperial College, London

MS. received 11*th January* 1951

ABSTRACT. The diffraction of plane waves by a curved linear lattice is studied by calculating the 'intensity' in reciprocal space. It is shown that the regions of high intensity are arc-shaped in character. Further, the intensity associated with a particular order of diffraction (considered as a function of the cylindrical radial coordinate in a plane parallel to that containing the curved lattice) can be represented by the square of a Bessel function of high index in the neighbourhood of the main peak, but dies off sharply on that side of the peak furthest from the origin. This 'cut-off' is characteristic of the total angle of bend as long as this angle of bend is less than 180°. The zero order diffraction peak is of a different type from the others, but is similar to the peak characteristic of a straight chain except for angles of bend greater than 180°.

§ 1

THE diffraction of plane waves by a curved linear lattice appears to have been considered only in the particular case in which the scattering points are arranged with uniform spacing on the circumference of a circle. This was done by Fock and Kolpinsky (1940) in connection with an investigation into the diffraction from a bent crystal. Their method will be described below. The more obvious case of an arc of scattering points, which is of greater interest to physicists, proved apparently intractable. A rather different approach has been attempted in this paper which does solve the problem, though the solution is restricted to the case of a reasonably large total number of scattering points.

The amplitude of waves scattered by a curved linear lattice is proportional (ignoring an 'obliquity' factor) to

$$G=\sum_{l} \exp\,[2\pi i\{\mathbf{r}_l\,.\,(\mathbf{s}-\mathbf{s}_0)/\lambda\}], \qquad \ldots\ldots(1)$$

where the amplitude is measured at a large distance from the lattice. The incident wave is assumed to be plane, $\mathbf{s}_0$ being a unit vector in the direction of incidence, $\mathbf{s}$ a unit vector in the scattering direction. The wavelength of the incident radiation is λ, and $\mathbf{r}_l$ a vector from an arbitrary origin to one of the scattering points. The vector $(\mathbf{s}-\mathbf{s}_0)/\lambda=\mathbf{H}$ is a vector in reciprocal space.

The centre of curvature is taken as origin, and the centre point of the arc as the zero of l (Figure 1). The total number of scattering points is $2L+1$. It follows that $|\mathbf{r}_l|=R$, where R is the radius of curvature. We can then write $(\mathbf{r}_l\,.\,\mathbf{H})=RH_p\cos(l\phi-\alpha)$, where $\phi=a/R$ and a is the distance between scattering points measured along the arc; H_p is the magnitude of the component of $\mathbf{H}$ in the plane of the arc, and the angles ϕ and α are measured from a line joining the origin to $l=0$ (the y axis).

Following the notation of Fock and Kolpinsky, we put $2\pi R=pa$ and $aH_p=\rho$ and then can write (1) in the form

$$G=\sum_{-l}^{+l} \exp\,[ip\rho\,(\cos 2\pi l_0/p-\alpha)]. \qquad \ldots\ldots(2)$$

The method used by these authors for a circle of scattering points consists in expanding the exponential in a series of Bessel functions

$$\exp(ip\rho\cos\theta)=J_0(p\rho)+2\sum_{q=1}^{\infty} i^q \cos q\theta J_q(p\rho).$$

Inserting this in (2),

$$G=\sum_l J_0(p\rho)+2\sum_l\sum_{q=1}^{\infty} i^q J_q(p\rho)\cos q(2\pi l/p-\alpha). \qquad \ldots\ldots(3)$$

The total number of points is p in this case; the summations in the second term of (3) can be interchanged, and then that over the cosine term will vanish except when $q=np$ (n being an integer), when the summation gives the value $p\cos np\alpha$. Hence

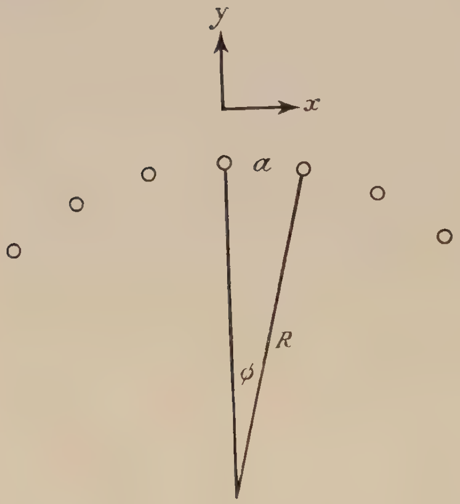


Figure 1.

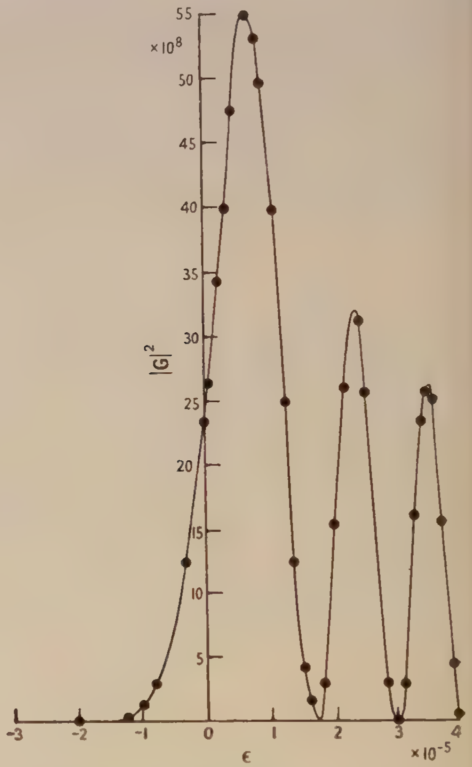


Figure 2. The intensity function $|\mathbf{G}|^2$ as a function of $\epsilon=1-\rho$ with $\beta=0$, for values of ϵ between -2×10^{-5} and $+4\times10^{-5}$. The curve continues up to $\epsilon=0{\cdot}015$.

$$G=pJ_0(p\rho)+2p\sum_{n=1}^{\infty} i^{np}J_n(p\rho)\cos np\alpha.$$

In discussing the value of $|\mathbf{G}|^2$, Fock and Kolpinsky average this over a range of α values at fixed ρ. This will eliminate the double cosine terms and yields

$$|\mathbf{G}|^2=p^2J_0^2(p\rho)+2p^2\sum_{n=1}^{\infty} J_{np}^2(p\rho). \qquad \ldots\ldots(4)$$

The function $J_{np}^2(p\rho)$ has a peak value near $\rho=n$, and if p is large each of the peak regions can be described by a single term of (4). The value of $|\mathbf{G}|^2$ in reciprocal space is a function of ρ only, and, considered as a function of ρ, consists of unsymmetrical peaks rising sharply on the side where $\rho<n$, and decaying slowly for $\rho>n$ apart from the characteristic oscillations of the Bessel function. This is illustrated in Figure 2, where the region $n=1$ has been calculated in another connection (§ 4).

§ 2

The above method does not lead to a correspondingly simple form when the circle is replaced by an arc. A different approach was tried, which consists in looking for regions of 'stationary phase' in the scattering functions. We expand the exponent in a term of (2) round a point $l=l_0$:

$$
\begin{aligned}
&ip\rho \cos (2\pi l/p-\alpha) = \\
&\qquad i\rho \cos (2\pi l_0/p-\alpha)[p-2\pi^2(l-l_0)^2/p+\dots] \\
&\qquad -i\rho \sin (2\pi l_0/p-\alpha)[2\pi(l-l_0)-4\pi^3(l-l_0)^3/3p^2+\dots]. \qquad \dots\dots(5)
\end{aligned}
$$

We choose l_0 such that

$$-\rho \sin (2\pi l_0/p-\alpha)=n, \qquad \dots\dots(6)$$

where n is an integer. We exclude the case $n=0$, which will be dealt with separately in § 4.

The first term involving n is $2\pi i n l$, which will give a factor of unity in G since l is an integer. It also follows from (6) that $\rho \cos (2\pi l_0/p-\alpha)=(\rho^2-n^2)^{1/2}$, and, expressing ρ as

$$\rho=n+\epsilon, \qquad \dots\dots(7)$$

this becomes

$$\rho \cos (2\pi l_0/p-\alpha)=(2n\epsilon+\epsilon^2)^{1/2}. \qquad \dots\dots(8)$$

Considering (5) again, the first of the remaining terms which contains l is quadratic in $(l-l_0)^2$. We can then write (5) as

$$-i[2\pi n l_0-p(2n\epsilon+\epsilon^2)^{1/2}+2\pi^2(2n\epsilon+\epsilon^2)^{1/2}(l-l_0)^2/p+4\pi^3 n(l-l_0)^3/3p^2+\dots]. \qquad \dots\dots(9)$$

It will be obvious that this has a stationary value for $l=l_0$. The method to be adopted in calculating is to replace the summation over l by an integration, which will be justified as the total number of points $2L+1$ is assumed to be large. It is to be expected that large values of $|\mathbf{G}|^2$ will be associated with small values of ϵ.

It further follows from (6) and (8) that

$$\cot (2\pi l_0/p-\alpha)=-(2n\epsilon+\epsilon^2)^{1/2}/n, \qquad \dots\dots(10)$$

where the positive sign of the square root is taken. It will be more convenient to replace the angle α by $\beta=\frac{1}{2}\pi-\alpha$; the direction $\beta=0$ will be parallel to the direction of the chain when the radius of curvature is infinite. We can hence write (10) in the form

$$\tan (2\pi l_0/p+\beta)=(2n\epsilon+\epsilon^2)^{1/2}/n \quad \text{or} \quad 2\pi l_0/p=-\beta+\tan^{-1}[(2n\epsilon+\epsilon^2)^{1/2}/n]. \quad \dots(11)$$

Now l_0 varies from $-L$ to $+L$; if $\epsilon=0$, β varies from $-2\pi L/p$ to $+2\pi L/p$, which represents the total angle of bend of the lattice. The region in reciprocal space for which (11) holds with $|l_0|<L$ will contain large values of $|\mathbf{G}|^2$; outside this there will be no regime of stationary phase and $|\mathbf{G}|^2$ will be expected to fall sharply.

It will be seen from (11) that there are really two possible values of l_0 for a fixed β and ϵ, depending on the two signs of the square root. We take the positive value and calculate G with this; the second position of stationary phase will automatically be found in the process of calculation. When $\epsilon > 0$ the limits of the arc-shaped region are given by $\pm \Psi$, where $\Psi = 2\pi L/p + \tan^{-1}[(2n\epsilon + \epsilon^2)^{1/2}/n]$. Though it may appear that the lower of these limits is inconsistent with (11), in actual fact both solutions $\pm(2n\epsilon + \epsilon^2)^{1/2}/n$ will appear in the calculations, as mentioned above. It will also be seen from (1) that G changes into its conjugate complex on inversion at the origin of reciprocal space. Hence $|\mathbf{G}|^2$ is symmetrical on inversion. Also the curved lattice is symmetrical on reflection in the y axis (Figure 1). If we combine these two properties, it follows that $|\mathbf{G}|^2$ is symmetrical on reflection in the x axis, and the arc-shaped region discussed above must extent to $-\Psi$ if it extends to $+\Psi$.

It remains to consider the case of $\epsilon < 0$ where the condition (8) cannot be fulfilled for real values of l_0. It is, however, still possible to use the same formal conditions for complex values of l_0, where this is treated as a parameter. One can then show that G can be written in the form

$$G = A\int_{-L}^{+L} dl \,.\, \exp[-2\pi^2(n^2-\rho^2)^{1/2}(l-l_0')^2/p]\exp[i(4\pi^3(l-l_0')^3/3p^2 + \dots)], \qquad \dots\dots(12)$$

where $A = \exp[p(n^2-\rho^2)^{1/2} + 2\pi n l_0'']$ and l_0', il_0'' are the real and imaginary parts of l_0.

The imaginary part of G will be zero as long as the value of the integral is independent of the limits. This will be the case as long as β is inside the critical arc by an amount which depends, of course, on L and p, but which is negligible for instance for $L = 10^6$, $p = 3{\cdot}6 \times 10^7$, a case considered below. The integration can be carried out numerically without difficulty. There is, however, a limiting case of some interest for values of $(n^2-\rho^2)^{1/2}$ which are not too small. Here the imaginary exponential in (12) can be replaced by unity and we obtain

$$G = [p/2\pi(n^2-\rho^2)^{1/2}]\exp[-p(n^2-\rho^2)^{3/2}/3n^2]. \qquad \dots\dots(13)$$

When p is large, the above can be shown to be identical with the Debye asymptotic formula for $pJ_n(p\rho)$ and $\rho < n$.

§ 3

The above identification of G with $pJ_{np}(p\rho)$ in a certain limiting case suggests that the relationship should hold over a wider field. It seems therefore advisable to see how far one can handle a Bessel function by the methods of §2. For simplicity we take p to be an even integer. In the Bessel function $J_p(p\rho)$, defined by the relation

$$J_p(p\rho) = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\exp[i(-p\theta + p\rho\sin\theta)]\,d\theta, \qquad \dots\dots(14)$$

we put $\theta = 2\pi l/p$, where l is now a continuous variable, then

$$pJ_p(p\rho) = \int_{-p/2}^{+p/2}\exp[i(-2\pi l + p\rho\sin 2\pi l/p)]\,dl.$$

If we expand, as before, round a point $l = l_0$ we can write

$$\begin{aligned} i(-2\pi l + p\rho\sin 2\pi l/p) = {} & -2\pi i l_0 - 2\pi i(l-l_0) \\ & + i\rho\cos 2\pi l_0/p[2\pi(l-l_0) - 4\pi^3(l-l_0)^3/3p^2 + \dots] \\ & + i\rho\sin 2\pi l_0/p[p - 2\pi^2(l-l_0)^2/p + \dots]. \end{aligned} \qquad \dots\dots(15)$$

Choosing l_0 such that $\rho \cos 2\pi l_0/p = 1$ ......(16)

it follows that $\rho \sin 2\pi l_0/p = (1-\rho^2)^{1/2}$, and (15) can be written as

$$i(-2\pi l_0 + p(1-\rho^2)^{1/2} - (1-\rho^2)^{1/2}\,2\pi^2(l-l_0)^2/p - 4\pi^3(l-l_0)^3/3p^2 + \ldots.),$$

which is identical with (9) for $n=1$ and $\rho = 1+\epsilon$. In fact the only difference in the calculation of G lies in the limits of integration. The above identification holds also for $\rho<1$ where a complex formulation for l_0 has to be used. Hence $G = pJ_p(p\rho)$ for $n=1$ as long as the value of G is independent of the limits of integration. This argument can be extended to higher values of n by replacing p by np in (14), leaving $p\rho$ unaltered. The case where $n=0$ is exceptional since the condition (16) then leads to $l_0 = \pm p/4$. This can be identified with G for $n=0$ only for total angles of bend greater than 180°. This point is discussed further in §4.

§4

The numerical calculation was carried out in a particular case in which $L = 10^6$, $p = 3{\cdot}6 \times 10^7$, i.e. a total angle of bend of 20°. These numbers are those chosen in a calculation by the author of the diffraction from a bent crystal and forms a limiting case to this. A single order $n=1$ is taken with $\beta = 0$, i.e. $|\mathbf{G}|^2$ is calculated in reciprocal space along the x axis.

As shown above, one expects a curve which should be identical with $p^2J_p^2(p\rho)$ in the neighbourhood of $\rho=1$ with certain restrictions. The numerical work checked the assumption that the limits of the integration are unimportant in the region shown in Figure 2, and also allows one to extend the calculation to cases where $\beta \neq 0$.

The curve shown agrees within 1% with a direct calculation of $p^2J_p^2$ at various points. The calculation shows the oscillatory character to be due to the phase linkage of the two regions of constant phase. When these are well separated, one can show that $|\mathbf{G}|^2 \simeq 4I^2(1+\cos\theta)$, where θ is the phase difference of the two regions and

$$I = \int_{-\infty}^{+\infty} \cos\,[(\rho^2-1)^{1/2}2\pi^2(l-l_0)^2/p]\,d(l-l_0).$$

This is a Fresnel integral and, putting in its value,

$$|\mathbf{G}|^2 \simeq \frac{p}{\pi}\frac{1}{(\rho^2-1)^{1/2}}(1+\cos\theta). \qquad \text{......(17)}$$

This is a good approximation to the values calculated from the first subsidiary maximum onwards.

The correspondence between $|\mathbf{G}|^2$ and $p^2J_p^2$ breaks down for large values of $(\rho^2-1)^{1/2}$ or $(2\epsilon)^{1/2}$. This can be seen if we look at the relation (11) between l_0 and ϵ with $\beta=0$. When l_0 reaches the value of L, the two positions of stationary phase have reached the edge of the lattice. The limiting value in the above calculation is $\epsilon = 0{\cdot}015$. In the neighbourhood of this value we should expect a sharper decrease in $|\mathbf{G}|^2$ with increasing ϵ.

We can give an approximate expression for $|\mathbf{G}|^2$ averaged over the subsidiary maxima; this is based on the same method as led to (17):

$$|\mathbf{G}|^2 \simeq 2(I_c^2 + I_s^2), \qquad \text{......(18)}$$

where $I_c = \int_{-\infty}^{L-l_0} \cos bx^2 dx$, $I_s = \int_{-\infty}^{L-l_0} \sin bx^2 dx$, and $b = 2\pi^2(2\epsilon+\epsilon^2)^{1/2}/p$.

A numerical examination of (18) shows that there is a marked change in $|\mathbf{G}|^2$ in a region of width $\Delta\epsilon$ approximately 0·001 around $\epsilon=0{\cdot}015$. At this latter point the value of $|\mathbf{G}|^2$ is one quarter of the value given by averaging (17); at $\epsilon=0{\cdot}016$, $|\mathbf{G}|^2$ has become negligible compared with the value given by (17). It is therefore justified to consider the curve as breaking off at the 'critical' value of ϵ.

The main difference when $\beta\neq0$ lies in the value of l_0 given by $2\pi l_0/p=\beta+\tan^{-1}(2\epsilon+\epsilon^2)^{1/2}$ which means that the two positions of stationary phase are symmetrical not about $l=0$ but about $p\beta/2\pi$. This has no effect on $|\mathbf{G}|^2$ (if $|\beta|<2\pi L/p$) until a value of ϵ is reached where l_0 approaches L. In this region, and beyond, equation (18) can be generalized without difficulty. In the immediate neighbourhood of $|\beta|=2\pi L/p$ the whole pattern does, however, change. In the above numerical case this will occur for $p\beta/2\pi\simeq9{\cdot}8\times10^5$. The detailed features will need the detailed curves on which Figure 2 is based. The general features can, however, all be seen from a study of the region, or regions, of stationary phase.

For values of n larger than unity the following differences will be found: the features near the main maximum will be those of $p^2J_{np}^{\ 2}(p\rho)$, and this means that the peak value will decrease as $n^{2/3}$ and will move out to larger values of ρ, the distance from $\rho=n$ being proportional to $n^{1/3}$; the initial rise will be less marked. These are features already noted by Fock and Kolpinsky for the circle.

The value of ϵ at which the intensity starts to drop off is given for $\beta=0$ by

$$2\pi L/p=\tan^{-1}[(2n\epsilon+\epsilon^2)^{1/2}/n]. \qquad \ldots\ldots(19)$$

If we solve for ϵ, we see that ϵ is proportional to n at least for low values of n, i.e. the 'width' of the region of high intensity increases with order roughly in proportion to the order.

The case of $n=0$ has to be handled separately. The diffraction peak differs from all others, not only in that it is symmetrical in reciprocal space, but that it can differ markedly from $p^2J_0^{\ 2}(p\rho)$ when the total angle of the bend is small.

If we refer to the main condition (7) which governed the previous considerations, i.e. $-\rho\sin(2\pi l_0/p-\alpha)=\rho\cos(2\pi l/p+\beta)=n$, it will be seen that if $\rho=\epsilon<1$ and $\beta=0$ it will be impossible to satisfy this relation for $n=0$ and still make $|l_0|$ less than L. In fact it is only possible to satisfy the relation at all under the above conditions if $l_0=\pm p/4$. This lies outside the actual number of points L for the angle of bend we have considered. We can, however, still use this value of l_0 as a parameter; the lth term in G then becomes

$$\exp\,[i\epsilon(p-2\pi(l-p/4)^2/p+\,\ldots\,.)].$$

In forming $|\mathbf{G}|^2$ the constant term $ip\epsilon$ will disappear. The remainder is then similar, for values of ϵ which are not too small, to that for higher orders when the 'critical' value of ϵ given by (19) is exceeded. The calculation can be performed numerically, but it will be sufficient for our purpose to give a rough value: $|\mathbf{G}|^2=2I_c^{\ 2}$, where $I_c=\int_0^L\cos\,[2\pi^2\epsilon(l-p/4)^2/p]dl$. This has the value $(2L)^2$ for $\epsilon=0$ and will be negligible compared with this when $\epsilon>4\times10^{-6}$ (for $L=10^6$, $p=3{\cdot}6\times10^7$).

For values of β different from zero (keeping $n=0$) the shape of $|\mathbf{G}|^2$ as a function of ϵ will vary markedly. This is due to the change in the parameter

$l_0=p/4-p\beta/2\pi$, which is reduced for instance to $l_0=0$ for $\beta=\pi/2$. In that case the value for G becomes

$$G=2\int_0^L \cos\,(2\pi^2\epsilon l^2/p+\ldots.)\,dl,$$

and here, for $\epsilon>10^{-5}$, $|\mathbf{G}|^2$ will vary as ϵ^{-1} in the above numerical case.

In this particular case of $n=0$ the diffraction peak shows also a marked change in character with the value of the total angle of bend $2\pi(2L+1)/p$. As soon as this becomes larger than 180° there will be regions of constant phase inside the limits of integration of G for all values of β. The character of the radial sections of $|\mathbf{G}|^2$ in reciprocal space is then similar to that of $p^2J_0^2(p\rho)$, as in the above-mentioned case of $\beta=\pi/2$, $n=0$ and values of ϵ which are not too small.

So far we have been concerned mainly with a definite value of the total angle of bend of the lattice. The variation of $|\mathbf{G}|^2$ with angle of bend can be deduced in most cases from the discussion given above; the transition to the case of a linear chain is, however, not quite straightforward, and is best described in terms which allow a more direct approach.

This consists in a direct expansion of (5), and the exponent becomes

$$[ip\rho\cos\alpha-i\rho\sin\alpha\,.\,2\pi l+i\rho\sin\alpha\,.\,4\pi^3l^3/3p^2-i\rho\cos\alpha\,.\,2\pi^2l^2/p+\ \ldots.].$$

The first term is a constant phase factor and can be ignored. The second term leads to the formula for the linear chain. The remaining terms characterize the effect of a finite radius of curvature. Taking $\alpha=\frac{1}{2}\pi$ $(\beta=0)$, the expression for G can be written as

$$G=\sum_l \exp\,[-i\rho 2\pi l+i\rho 4\pi^3l^3/3p^2+\ \ldots.].$$

When the largest term in l^3 is small compared with π, one would expect $|\mathbf{G}|^2$ to approximate closely to the form for the linear chain. We take as a criterion

$$4\pi^3nL^3/3p^2=\pi/10, \qquad \ldots\ldots(20)$$

where $\rho=n$ refers to the nth order. This gives, in the above numerical case of $L=10^6$, a limiting value of $p=6\times10^9$, or a total angle of bend of approximately 0·1° for $n=1$. Equation (20) also shows that this limiting value of p is proportional to $n^{1/2}$.

The transition to the linear chain has been studied numerically for a chain containing a small number of scattering centres. This investigation shows that (20) is a reasonable criterion. The detailed results will be reported at a later date.

REFERENCE

Fock, V. A., and Kolpinsky, A., 1940, *J. Phys. U.S.S.R.*, **3**, 125.

The Morphology of Zinc Oxide Smoke Particles

By J. M. COWLEY, A. L. G. REES and J. A. SPINK *

Chemical Physics Section, Division of Industrial Chemistry, Commonwealth Scientific and Industrial Research Organization, Melbourne, Australia

MS. received 21*st December* 1950

ABSTRACT. Single-crystal electron diffraction patterns with doubled spots obtained from zinc oxide smoke particles originate from pairs of thin sheets which are commonly seen in electron micrographs of this material joining up the members of a group of four spines. The plane of each sheet is parallel to a $(10\bar{1}0)$-type lattice plane. The sheets are continuous with one of the spines and spaced at 120° intervals around this spine. The other three spines, and additional spines which grow from corners of the sheets, are related to this one spine and the sheets by twinning on $(11\bar{2}2)$ planes. Intensity measurements of single-crystal patterns allow the shape-transform and thickness of the sheets to be deduced. Sharp lines and extra spots indicate the occasional presence of lattice imperfections in the form of parallel faults. Electron micrographs show wide deviations from the idealized group of spines and sheets.

§ 1. INTRODUCTION

Electron diffraction patterns of zinc oxide smoke, taken under conditions of high resolution, show considerable fine structure. Diffraction spots are elongated or split into several components. These effects have been correlated with the needle-like shape of the zinc oxide crystals. Electron microscopic examination has shown that zinc oxide smoke consists largely of groups of four spines, each 50 to 500 A. in diameter and up to 5 μ long, joined at a common centre in approximately tetrahedral array. Each spine consists of a single crystal of zinc oxide elongated along the *c*-axis. Fuller (1944) examined these 'fourlings' by stereoscopic electron microscopy and deduced that the spines are related by twinning on $(11\bar{2}2)$ planes with a slight distortion at the twinning faces, so that, of the six angles between pairs of spines, two are 97·6° and four are 116·1°. Electron diffraction spots from such spines are extended by the effect of refraction of beams passing through non-parallel faces and effects arising from the small dimensions of the crystals across the spines. Such effects have been considered in detail elsewhere (Rees and Spink 1950 and paper in preparation).

In addition to the elongated spots given by the spines, diffraction patterns contain sharp spots which are often of relatively high intensity and sometimes markedly separated from the rings. Such sharp spots frequently occur in pairs or triplets. In these cases the assumption that the separation of the components of the pairs or triplets is a result of refraction by a hexagonal prism often gives a very high apparent value for the inner potential.

These sharp spots are apparently given by the thin sheets of zinc oxide which often appear in electron micrographs of the smoke, associated with groups of spines (Figures 7 and 8, Plate II †). The thickness of these sheets is approximately the same as, or slightly less than, the diameters of the spines. The sheets appear to stretch between spines of the fourlings.

---

* Division of Tribophysics, C.S.I.R.O.

† For Plates see end of issue.

Examination of many micrographs leads to the conclusion that they form a webbing, with each sheet lying in the plane of two spines, rather than a hollow tetrahedral bag with the spines growing from the corners. Information on the structure of these sheets and their relation to the spines, and on the origin of the sharp pairs and triplets of diffraction spots has been obtained from single-crystal diffraction patterns given by zinc oxide smoke crystals.

§2 SINGLE-CRYSTAL ELECTRON DIFFRACTION PATTERNS

Electron diffraction patterns were obtained from zinc oxide smoke using a high-resolution diffraction camera (Cowley and Rees, unpublished). A narrow beam, defined by the small electron source and a $20\,\mu$ aperture, was focused at the specimen level so that a very small area of the specimen was exposed to high intensity. Individual crystals then gave strong reflections. In some cases strong cross-grating patterns were observed. The individual spots were sharp, and showed no elongation or multiplicity due to refraction or small crystal dimensions. Hence these patterns were attributed to large sheets of zinc oxide. The most commonly occurring cross-grating consisted of $h0\bar{h}l$ spots.

Several of the cross-grating patterns recorded show a doubling of the spots. Both components of a doubled spot are sharp. Their separation is zero on the Laue circle of strong reflections and increases with distance from the Laue circle. Figure 1 (Plate I) is a cross-grating of $h0\bar{h}l$ reflections showing such doubling; for the sake of clarity a background of weak rings and odd spots has been removed by masking. For each pair of spots, the separation is perpendicular to the line of $000l$ spots, and parallel to the line of $h0\bar{h}0$ reflections. 'Forbidden' reflections, such as the 0001, 0003 and $10\bar{1}3$, are present with considerable intensity and show the same doubling. The $10\bar{1}3$ reflection is on the Laue circle and is particularly strong. The two components of a doubled spot are equally displaced from the points of the rectangular cross-grating net defined by the non-doubled spots on the Laue circle.

In similar cross-grating patterns for which the Laue circles are small, the doubling is visible only outside the circles. It is evident that such a doubling of the cross-grating may be interpreted in terms of the intersection of the Ewald sphere of reflection with reciprocal-lattice points drawn out into two sharp spikes inclined at equal angles to the plane of $h0\bar{h}l$ reflections and both perpendicular to the *c*-axis. Thus Figure 2 shows a section of the Ewald sphere cutting through such pairs of spikes along a row of reflections perpendicular to the *c*-axis.

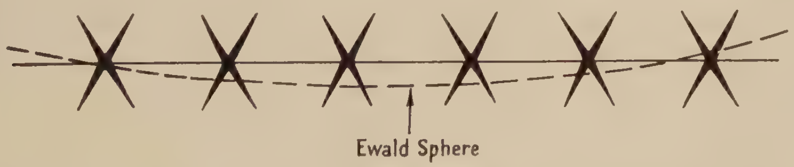


Figure 2. Section of the reciprocal lattice showing how the doubling of the cross-grating spots arises from the intersection of the Ewald sphere with reciprocal-lattice points drawn out into two sharp spikes.

The sharpness of the components of the doubled spots precludes the possibility that the twin spikes through the reciprocal-lattice points could be given by the shape-transform of the polygonal cross section of a spine-like crystal. Rather, each spike appears to be defined by the shape-transform of a very thin plate-like crystal. The observed pattern could then be given by two sheets with identical lattice orientation, but inclined at equal and opposite angles to the electron beam and meeting along a line parallel to the *c*-axis.

The angle between the spikes in the reciprocal lattice may be calculated readily from the separations of the components of pairs of spots and their distances from the Laue circle. From 41 well-resolved pairs of spots in the pattern of Figure 1, the angle between spikes was calculated as $58{\cdot}2 \pm 3{\cdot}7°$, and from 40 pairs in another pattern $60{\cdot}6 \pm 6{\cdot}9°$. Thus the zinc oxide sheets concerned are each inclined at about 60° to the electron beam, so that the angle between the sheets is either about 60° or 120°. For a 60° angle the beam would pass through the two sheets in succession. Diffracted beams from the first sheet would be re-diffracted, giving rise to extra spots around some pairs. Such secondary diffraction effects have been observed, for example, in paraffins, where there is a superposition of lamellar crystals of scattering power comparable with that of the zinc oxide sheets (Cowley, Rees and Spink 1951). However, no extra spots are detectable and hence the angle of 120° seems more probable.

It is to be expected that the large bounding faces of the zinc oxide sheets would be parallel to principal planes with low indices and high atom density. All observed effects may be explained by assuming the sheets to be bounded by $(10\bar{1}0)$ planes, which are the planes of second highest atom density. If the sheets are bounded by $(10\bar{1}0)$-type planes, and are joined along a line parallel to the *c*-axis, the angle between a pair of sheets is 120°, and a cross-grating of $h0\bar{h}l$ reflections is obtained when the electron beam is perpendicular to the *c*-axis and inclined at 60° to each sheet.

Electron micrographs show that the zinc oxide sheets stretch between pairs of spines. For 'fourlings' with no sheets present, Fuller showed that the spines have a symmetrical arrangement produced by $(11\bar{2}2)$-twinning. Pairs of sheets joining the spines of such fourlings would be inclined to one another at $113{\cdot}5°$ or $123{\cdot}5°$, in which case the sheets either would not have identical lattice orientation, or would not be bounded by low-index lattice planes.

The electron diffraction single-crystal patterns thus indicate that when a pair of sheets is present, the sheets do not join up spines arranged in the form of such a fourling as Fuller found, but rather, the two sheets form continuations of the lattice of one spine, with an angle of 120° between them.

§3. SHAPE-TRANSFORM AND THICKNESS OF SHEETS

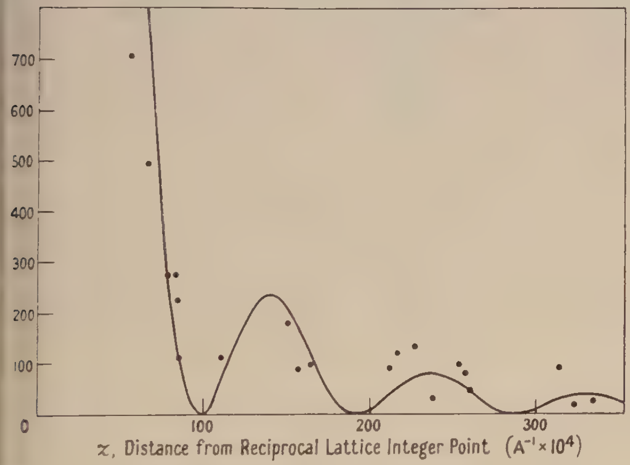
For a thin sheet of uniform thickness and large lateral dimensions, the scattering point on the reciprocal lattice corresponding to each reflection is drawn out into a thin spike, perpendicular to the sheet. The shape-transform defining the variation of scattering power along the spike is given by the kinematic theory of electron diffraction as $I(z)/I_0 = (\sin^2 \pi N d z)/(\pi d z)^2$, where N is the number of scattering points in the direction perpendicular to the sheet, d is the distance between them, and z is the reciprocal lattice coordinate. MacGillavry (1940) has shown that the dynamic theory of electron diffraction gives, in effect, a shape-transform different from this, but approaching it for weak reflections or very thin sheets.

From the doubled cross-gratings it is possible to deduce the shape-transforms of the sheets by comparing the observed intensity of a reflection with the theoretical integrated intensity. The observed intensity is proportional to the scattering power where the Ewald sphere cuts the spike. The distance z of this point of intersection from the Bragg reflection position is determined from the separation of the components of pairs of spots.

Thus $S(z)=KI_{h0\bar{h}l}(z)/E^2_{h0\bar{h}l}$, where K is a constant and $E_{h0\bar{h}l}$ is the electron diffraction structure factor for the $h0\bar{h}l$ reflection.

The estimation of the intensity terms $I(z)$ is rendered uncertain by the wide range of intensities present in the cross-grating pattern and also by the effect of dynamic interaction or secondary scattering. It is apparent that the intensities may be modified considerably by this latter effect, since forbidden reflections such as 0001, 0003 and $10\bar{1}3$ occur quite strongly. However, even without correction for these effects, the intensities may be used to indicate the general form and linear dimensions of the distribution of scattering power.

From patterns similar to Figure 1, the intensities of the weaker spots were estimated visually. For the stronger spots, it was assumed that the intensity is proportional to the square of the corrected half-widths. The values of $S(z)$ found from the pattern of Figure 1 are plotted against the distance along the spike z in Figure 3. For comparison, the curve of $(b/a^2z^2)\sin^2 az$ is drawn with the parameters a and b adjusted to give the best fit of the experimental points. The value of a used corresponds to a thickness of the sheet equal to 104 A.



ure 3. Shape transform of a zinc oxide sheet. The experimental points are the values of $S(z)$, the ratio of observed to theoretical intensity, plotted against z, the distance from a reciprocal-lattice integer point. The full curve is the theoretical (kinematic) shape-transform.

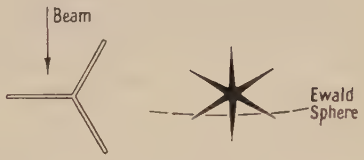


Figure 5. Origin of the triplet shown in Figure 4. The beam is perpendicular to one of three symmetrically arranged sheets. The intersection of the Ewald sphere with the three corresponding spikes about the $11\bar{2}0$ reciprocal–lattice point gives three spots on a radial line.

§ 4. FURTHER ELECTRON DIFFRACTION OBSERVATIONS

Doubling of the spots was observed in several cross-grating patterns other than the $h0\bar{h}l$ patterns illustrated. In addition, most ring patterns from zinc oxide smoke showed a number of pairs of sharp spots given by pairs of sheets not oriented with a principal axis nearly parallel to the beam. Two such pairs are visible in Figure 4 (Plate I). Also, in several cases triplets of sharp spots were observed. A radial group of three spots is apparent on the $11\bar{2}0$ ring in Figure 4. There is a similar group on the same ring diametrically opposite that shown. The spots of these groups are too sharp to be given by a prismatic crystal. Such groups of three spots could be produced by diffraction from a group of three sheets with identical lattice orientation but with planes 120° apart. With the electron beam almost perpendicular to one sheet, the triplets of spots would arise as shown in Figure 5. Assuming the three sheets to be of equal area, the path length of the beam, and so the loss of intensity due to inelastic scattering, would be less in the sheet perpendicular to the beam, so that the central spot should be more intense than the other two. This is as observed.

Figure 6 (*a*) is a sketch of a cross-grating pattern showing further new features.

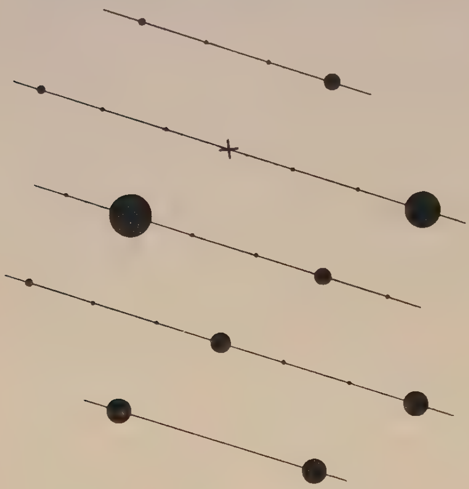


Figure 6 (*a*). Sketch of a cross-grating pattern with sharp lines and extra spots between normal diffraction spots.

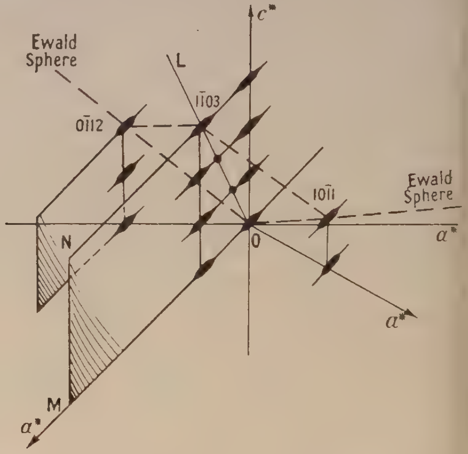


Figure 6 (*b*). Diagram of the reciprocal lattice showing the origin of the sharp line and spots between the origin and the $1\bar{1}03$ reflection in Figure 6(*a*). The Ewald sphere, considered plane (dashed lines) cuts the plane of scattering power, OMN, in the line OL, and the extended spikes through the lattice points in sharp spots.

These features were rather too weak in the original pattern for satisfactory reproduction. The cross-grating spots are joined by a series of sharp parallel lines. Weak sharp spots occur on these lines at positions one-third and two-thirds of the distance between cross-grating spots. These effects may be explained on the assumptions (i) that the sheet giving the cross-grating is very thin, so that the reciprocal lattice spikes are extended and there is appreciable scattering power all the way along the line joining two reciprocal-lattice points, and (ii) that the sheet is traversed by a series of faults perpendicular to the *c*-axis and therefore perpendicular to $(11\bar{2}0)$ planes. These faults would divide parts of the sheet into long needle-like prisms, which would give planes of scattering power in the reciprocal-lattice space. The intersection of the Ewald sphere, assumed plane, with such a system of planes and extended spikes would give the observed effects, as can be seen from Figure 6 (*b*).

In the same way it is possible to explain the occurrence of the weak, sharp lines sometimes observed joining the spots of $h0\bar{h}l$ cross-gratings.

§ 5. ELECTRON MICROSCOPE OBSERVATIONS

Electron micrographs such as Figures 7 and 8 (Plate II) lend support to the conclusions reached from the electron diffraction patterns. In each micrograph the central group seems to consist of four spines meeting at a point and three well-developed sheets. The sheets appear to be firmly attached to, and continuous with, the spine which points to the top of the picture in each case. Their connection with the other three spines seems less firm. A further feature shown in the micrograph is the occurrence of spines starting not from the centre of a fourling but from corners of the sheets remote from the centre. In this way as many as seven spines may be associated with the three sheets. In both Figures 7 and 8 six spines may be seen in the central group.

In the original negative of Figure 8, a strong 'dark-field' diffracted image is clearly visible. The outline of this image has been dotted in the reproduction. This diffracted image includes most of the vertical spine and the two well-exposed sheets. This is a clear indication that the crystal lattice in the two sheets has the same orientation as that of the spine.

The conclusion of Fuller that $(11\bar{2}2)$-twinning is the most probable relationship between zinc oxide spines may be applied to the present case. If the three bottom spines are related to the vertical one by such twinning, they will be coplanar with the sheets, and inclined at an angle of 116° to the vertical spine. If the spines which start from the outer corners of the sheets are related to the lattice of the sheets by $(11\bar{2}2)$-twinning, they will make an angle of 64° with the vertical spine. The idealized group of seven spines and three sheets is then as shown in Figure 9.

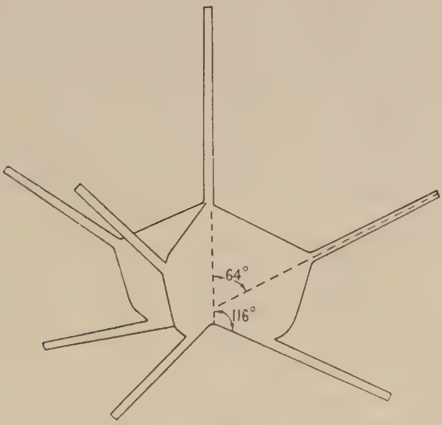


Figure 9. Idealized group of three sheets and seven spines, related by $(11\bar{2}2)$-twinning.

Accurate stereo-micrograph measurements of the angles between spines in the presence of sheets have not been made. However, measurements made on suitable oriented sheets in normal micrographs give an indication that the angles are as predicted. In Figure 7 the largest sheet appears to be almost perpendicular to the electron beam. The angles between the images of the vertical spine and the other two spines in the plane of this sheet are 66° and 117°. In Figure 8 the two well-exposed sheets appear to be almost equally inclined to the beam. Assuming that they are both at 60° to the beam, the angles between the spines are found to be 65° and 115°.

The same general relationship between sheets and spines were found to apply for a very complex group. This group includes one exceptionally large sheet of over 5μ extent, and two examples of a long narrow sheet stretching between two almost parallel spines. The sheets and spines of the group are bent by contact with a supporting film of collodion. They are traversed by very well-defined extinction contours corresponding to regions giving Bragg reflections. Such contours have previously been observed in micrographs of thin plates of mica (Heidenreich and Sturkey 1945), single crystals of metals (Heidenreich 1949) and other extensive crystals.

§ 6. DISCUSSION

The combined evidence of electron diffraction and electron microscopy indicates that, when two or three sheets are formed with a group of spines in zinc oxide smoke, they often grow so that the crystal lattice in all sheets is continuous with that of one of the spines. The sheets are bounded by $(10\bar{1}0)$ planes and spaced at 120° intervals about the spine. In the idealized group of four concurrent spines the other three spines are related to this spine and the sheets by twinning on $(11\bar{2}2)$ planes. The sheets are usually bounded by straight edges. Corners of the sheets, remote from the centre of the group, may act as nuclei for the growth of further spines with a $(11\bar{2}2)$-plane twinning relationship with the lattice of the sheet.

Many deviations from this idealized structure occur. Frequently only one sheet is well developed and often some spines are missing. Some samples of zinc oxide smoke contain very few sheets. Others contain a large proportion of very complex groups with many spines and sheets radiating from a large central nucleus. There is some evidence that occasional sheets may be bounded by other than $(10\bar{1}0)$ planes. In one doubled $h0\bar{h}l$ cross-grating the stronger spots were given by a sheet at 60° to the beam, but weaker spots from a crystal with almost identical lattice orientation showed no displacement from the points of the rectangular net. These weaker spots must have been given by a sheet perpendicular to the beam, inclined at an angle of 150° to the other sheet, and so bounded by $(11\bar{2}0)$ faces.

The sheets and spines may be distorted or imperfect; they may be bent elastically by contact with other particles or with a supporting film. The occurrence of continuous lines and extra spots in cross-grating patterns indicates inelastic distortion in the form of faulting.

ACKNOWLEDGMENT

The authors wish to thank Mr. J. L. Farrant for taking the electron micrographs of the zinc oxide specimens.

REFERENCES

Cowley, J. M., Rees, A. L. G., and Spink, J. A., 1951, *Proc. Phys. Soc.* A, **64**, 609.
Fuller, M. L., 1944, *J. Appl. Phys.*, **15**, 164.
Heidenreich, R. D., 1949, *J. Appl. Phys.*, **20**, 993.
Heidenreich, R. D., and Sturkey, L., 1945, *J. Appl. Phys.*, **16**, 97.
MacGillavry, C. H., 1940, *Physica*, **7**, 329.
Rees, A. L. G., and Spink, J. A., 1950, *Acta Crystallogr.*, **3**, 316.

The Coagulation and Deposition of a Stirred Aerosol

By J. CORNER and E. D. PENDLEBURY
Ministry of Supply, Fort Halstead, Kent

MS. received 29*th January* 1951, *and in amended form* 12*th March* 1951

ABSTRACT. Langstroth and Gillespie have determined experimentally the rate of change of particle number in ammonium chloride smokes in still and moving air, and have analysed their results into deposition on the walls and coagulation inside the aerosol cloud. It is shown here that the empirical constants describing these processes can be explained on theoretical grounds, both in order of magnitude and as regards their variation with the rate of stirring.

§ 1. INTRODUCTION

LANGSTROTH and Gillespie (1947) have observed the change of particle size and number in an ammonium chloride smoke in a chamber of size about one cubic metre, for various intensities of stirring. They have shown (cf. also Langstroth, Gillespie and Pearce 1949) that their results can be reproduced by assuming certain expressions for the loss of particles to the wall and for the rate of coagulation in the cloud. The empirical parameters describing such processes were obtained as functions of the intensity of stirring.

It is of interest to see whether these empirical results can be explained theoretically. This would enable the results to be generalized to other conditions of particle diameter, chamber size and turbulence.

It will appear in the present paper that the theoretical expectations do agree with the experimental results to within the uncertainty left by the method of stirring, which generates a fluid motion not closely resembling any whose turbulence has been studied.

§ 2. COAGULATION IN THE AEROSOL

The theoretical rate of coagulation of a substantially monodisperse aerosol was obtained by Smoluchowski (1917), both for a stationary medium and for one flowing in a laminar fashion. It was assumed that l, the mean free path of air molecules, was small compared with the particle radius r, and that the latter was large compared with the mean free path of the particles. The correction needed when the second of these assumptions is not satisfied has been obtained by Fuchs (1934); in the experiments of Langstroth and Gillespie the correction to the Smoluchowski law is of order 1%, and will be neglected here. Patterson, Whytlaw-Gray and Cawood (1927) have based a correction for finite values of l/r on the experimental mobility of oil drops in air (Millikan 1923). The Smoluchowski result for stationary medium is, with this correction,

$$-\frac{1}{n^2}\frac{dn}{dt}=K_0=4\boldsymbol{k}T(1+Al/r)/3\eta, \qquad \ldots\ldots(1)$$

where n is the number of aerosol particles per unit volume, η is the viscosity and T the temperature of the medium, and $\boldsymbol{k}$ is Boltzmann's constant. For the Langstroth and Gillespie experiments, r is of order 5×10^{-5} cm., that is, roughly $5l$; in this region A is approximately 0·9. Although the correction term is therefore not negligible, it does vary sufficiently slowly during any one experiment to be given a constant value for each case.

In the Smoluchowski treatment the molecular forces appear only in the assumption that particles stick together after a collision. The long-range effects of van der Waals forces have been taken into account by Tikhomirov, Tunitzki and Petrjanov (1942), who have shown that a new factor, independent of particle size and of order 1·2, appears in K_0. They have shown also that the enhanced coagulation rate found by Smoluchowski for particles of unequal sizes averages to a correction of only a few per cent for reasonable size distributions.

In the presence of a laminar flow with a velocity gradient $\partial u/\partial x$ normal to the stream lines, Smoluchowski finds the additional contribution to the coagulation rate

$$-\frac{1}{n^2}\frac{dn}{dt}=K_1=\frac{32}{3}r^3\left|\frac{\partial u}{\partial z}\right|. \qquad \ldots\ldots(2)$$

Langstroth and Gillespie have shown that (1) represents their coagulation rate in still air and reasonable corrections for size and distribution and van der Waals forces improve the agreement; thus it remains only to show that the variation of coagulation rate with the intensity of stirring is reproduced by the term K_1, provided that we can show that the flow is laminar over the distances, of the order of 10^{-2} cm., over which the particles have to diffuse for coagulation.

The flow inside the chamber used by Langstroth and Gillespie was rather complicated. Inside the cube, of side 112 cm., a paddle 60·5 cm. wide rotated about a horizontal axis. The paddle carried two vanes, 32·5 cm. × 60·5 cm., pierced by 15 holes 5 cm. × 5 cm., the whole being symmetrical about the axis. The vane could be oscillated through 180° from the vertical at any desired frequency, and the intensity of air motion $\bar{u}$ was obtained by averaging the reading of a non-directional thermocouple anemometer over various positions in the chamber. Air speeds ranged up to 80 cm/sec.

It has been pointed out to us by a referee that if the torque on the paddle spindle had been measured it would have been possible to deduce the velocity gradient, effective in coagulation, from the rate of dissipation of energy in the fluid. Unfortunately the torque was not measured by Langstroth and Gillespie, and it is therefore necessary to use a less direct method, introducing parameters which can be estimated roughly by comparison with more familiar turbulent flows. Let λ be the microscale of turbulence (Taylor 1935a) and u' the root mean square of the difference of some velocity component from its steady value. Then if we can show that λ is greater than 10^{-2} cm. we can take the flow to be laminar over the distance the aerosol particles need to diffuse, and may use (2) with

$$|du/dz|\simeq u'/\lambda. \qquad \ldots\ldots(3)$$

The values of λ and u' have unfortunately never been determined for this arrangement of vanes oscillating in a cubical box. A very rough estimate may be made by idealizing the paddle to a mesh, such as is used in a wind tunnel. The mesh size M for this paddle is of the order of 10 cm. We have (Taylor 1935b) $\lambda/M\simeq 2(\nu/Mu')^{1/2}$ where ν is the kinematic viscosity of the fluid. For air, with $M=10$ cm. and equating u' to its upper bound $\bar{u}$, $\lambda\sim 2{\cdot}5\,\bar{u}^{-1/2}$ when λ is in cm. and $\bar{u}$ in cm/sec. The largest value of $\bar{u}$ in the experiments of Langstroth and Gillespie was 100 cm/sec. Thus λ is indeed always greater, by at least one order of magnitude, than the distance over which the aerosol particles must diffuse, and, using (3), the velocity gradient is $|du/dz|\sim 0{\cdot}4(u')^{3/2}$.

It remains to estimate u' from the known frequency of the paddle and the corresponding value of the intensity of motion $\bar{u}$. These two quantities were shown experimentally to be proportional. The flow at 90° from a vane, an average condition, is about 5 mesh-lengths downstream, where in a wind tunnel u' is about 10% of the main stream velocity (Taylor 1935 b). However, it seems likely that because of the reversal of the paddle it would be better to idealize the flow to that behind a mesh in a wind tunnel when there is already a strong turbulence in the flow in front of the mesh. The fluctuation might then be of the order of 30% of the 'main-stream velocity'. The latter is itself not very easy to interpret in the experimental arrangement used by Langstroth and Gillespie, but it is obviously greater than $\bar{u}$; for example, the velocity of the end of the paddle is three times as large as $\bar{u}$. Thus we may expect u' to be rather less than $\bar{u}$, but not by any large factor. If we write $u' = h\bar{u}$ we can determine h by equating theoretical and experimental values of the additional coagulation rate due to turbulence. For the experiments with $\bar{u}$ equal to 50 m/min., the particle radius does not alter greatly with time, and is about half a micron (Figure 3, Langstroth and Gillespie 1947), while the coagulation rate due to stirring is 3×10^{-8} cm^3 min^{-1}. We find $h = 1{\cdot}2$. A value a little less than unity was to be expected, but the discrepancy is small and could be attributed to the necessarily rather rough estimate of λ, the microscale of turbulence, or to an error in our choice of the mean radius.

For a series of aerosols the coagulation rate due to turbulence would be proportional to $\bar{u}^{3/2}$ if the particle radius were the same in all cases. Even if the smokes had initially the same size distribution the different intensities of coagulation would set up differences in size. The effect would, however, be small if loss to the walls were prominent, as in the experiments of Langstroth and Gillespie. It is found, indeed, that the $\bar{u}^{3/2}$ law fits their results quite well.

It is interesting to compare the turbulence in the experiments of Langstroth and Gillespie with that deduced by Tunitzki (1946) to be appropriate to the coagulation of aerosols in the atmosphere. Tunitzki found du/dz to be of the order of 10 sec^{-1}. This would be attained, in the experiments of Langstroth and Gillespie, at $\bar{u}$ equal to about 5 m/min., or a tenth of their maximum rate of stirring.

3. DEPOSITION ON THE WALLS FROM THE UNSTIRRED AEROSOL

We investigate next the loss due to deposition on the walls of the containing vessel. There are three mechanisms of deposition, namely sedimentation, molecular diffusion and eddy diffusion. We consider first the experiments in which there is no turbulence and so no eddy diffusion.

Over the vertical faces of the cube deposition is due to molecular diffusion alone, but the deposit by this mechanism is very small and accounts at the most for only 3% of the total amount deposited. The deposition over the top will be less than that over the sides, for sedimentation will oppose the diffusion mechanism. We may therefore neglect this as a mechanism of deposition from the non-turbulent aerosol. Over the base of the cube we get deposition due to sedimentation and molecular diffusion; the equation describing this process is, neglecting the effects of the vertical walls of the cube,

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} + V\frac{\partial c}{\partial x}, \qquad \ldots\ldots(4)$$

where the axis of x is taken vertically upwards, $x=0$ being the base of the cube, and c is the local mass concentration of the particles. The coefficient of molecular diffusion is $D=(1+Al/r)\boldsymbol{k}T/6\pi\eta r$, where r is the radius of the particle. V is the sedimentation velocity of the particles.

Within the aerosol convection currents will be set up as a result of small temperature variations. These currents, though small, will cause the mass concentration to tend to be uniformly distributed throughout the volume. It will be noted that such currents also cause some of the particles to be deposited by the mechanism of eddy diffusion but it is clear, from the observed variation of deposition with the rate of stirring, that this effect will be negligible.

On the assumption that the mass concentration is uniform within the cube, except very close to the base, the rate of deposition can be derived from equation (4). Suppose that at all heights greater than a value ϵ, small compared with the depth of the cube, the concentration is uniform, and therefore equal to the mean value $\bar{c}$, this condition being maintained by thermal currents. We can identify $\bar{c}$ with the experimental mass concentration m. Equation (4) then applies to $0 \leqslant x \leqslant \epsilon$. We separate the variables by assuming a solution of the form

$$c \equiv X(x)\, T(t). \qquad \ldots\ldots(5)$$

This implies that the shape of the concentration–distance curve is independent of time. This is not normally the case in mathematical diffusion problems, where the usual condition is that c is initially uniform. In the Langstroth–Gillespie experiments, however, a small fan was operated in the chamber for five minutes before observations began, and it is reasonable to suppose that this stirring would set up the steady-state shape for the concentration of the 'concentration boundary layer' of thickness ϵ. We use (5), therefore, and hence

$$\frac{1}{T}\frac{dT}{dt} = -\alpha \qquad \ldots\ldots(6) \quad \text{and} \quad D\frac{d^2X}{dx^2} + V\frac{dX}{dx} = -\alpha X, \qquad \ldots\ldots(7)$$

where α is a constant whose value is decided by the conditions at the wall.

From (5) and (6), $\partial c/\partial t = -\alpha c$ and in particular, putting $x=\epsilon$,

$$d\bar{c}/dt = -\alpha\bar{c} \qquad \ldots\ldots(8)$$

while from (7), $X=e^{-Vx/2D}[A\exp\{(V^2+4\alpha D)^{1/2}x/2D\}+B\exp\{-(V^2+4\alpha D)^{1/2}x/2D\}]$ We assume that $4\alpha D/V^2 \ll 1$, leading to $X=Ae^{-\alpha x/V}+Be^{-Vx/D}$. When $x=0$, $c=0$, so that $X=A[e^{-\alpha x/V}-e^{Vx/D}]$. Let the length of the side of the cube which contains the aerosol be L. By considering the rate of loss to the base we obtain

$$d\bar{c}/dt = -(D/L)(\partial c/\partial x)_{x=0} = -VAT(t)/L.$$

Now $\bar{c}=AT(t)[e^{-\alpha\epsilon/V}-e^{-V\epsilon/D}] \simeq AT(t)e^{-\alpha\epsilon/V}$ if we neglect $e^{-V\epsilon/D}$. Hence $d\bar{c}/dt = -V\bar{c}e^{\alpha\epsilon/V}/L$. Comparison with equation (8) gives $\alpha = Ve^{\alpha\epsilon/V}/L$. We shall show that $e^{-\alpha\epsilon/V} \simeq 1$, giving $\alpha = V/L$. V is obtained from Stokes' law,

$$V = 2\rho g r^2(1+Al/r)/9\eta \qquad \ldots\ldots(9)$$

where ρ is the density of each particle. Typical conditions in the experiments of Langstroth and Gillespie would correspond to particles of radius $0{\cdot}6\,\mu$ and density 1 gm/cm^3, suspended in air at room temperature, giving $V=0{\cdot}37$ cm. min^{-1}. L was 112 cm., so that $\alpha = 3{\cdot}3 \times 10^{-3}$ min^{-1}. This is certainly of the same order as the observed α. A more exact agreement is not to be expected, since the effective density of the particles is doubtful, as is also the average radius to be used. Actually the constants used here were chosen to give the observed α exactly. We can conclude that sedimentation is adequate to explain the deposition observed by Langstroth and Gillespie in unstirred aerosols.

We now verify that $V^2 \gg 4\alpha D$, that $e^{-\alpha\varepsilon/V} \simeq 1$, and that $e^{-V\varepsilon/D}$ is negligible. Since D is of order $2{\cdot}5 \times 10^{-7}\,\mathrm{cm}^2\,\mathrm{sec}^{-1}$, V^2 is greater than $4\alpha D$ by a factor of order 10^6. Also α/V is of order 10^{-2} and ϵ will at most be of order 10 cm., so that the difference between $e^{-\alpha\varepsilon/V}$ and unity is not more than 10% in the experimental conditions of Langstroth and Gillespie. Also $V\epsilon/D$ is of order $10^4\epsilon$, when ϵ is measured in centimetres, so that $e^{-V\varepsilon/D}$ is certainly negligible.

§4. DEPOSITION FROM THE TURBULENT AEROSOL

We are concerned principally with the conditions near the boundary of the containing vessel, where the mean motion will be parallel to the sides. We postulate that there is a layer of thickness δ adjacent to each side in which the mean motion is parallel to the side, and that outside this layer the concentration is everywhere $\bar{c}$, a function of time. The fluid motion within the layer is turbulent and consequently there are random fluctuations in the motion, the magnitude of which tends to zero as the surface is approached. The situation may be likened to the flow near the walls of a pipe when the turbulent regime has extended over the whole cross section, that is, after the 'inlet length' has been passed. The fluctuations near the wall give rise to eddy diffusion, the coefficient of eddy diffusion being $D_e = K_e x^2$, where x is the distance from the side, and $K_e = K_0^2 d\bar{u}\; dx$ where K_0 is the Kármán turbulence constant, which for fully turbulent flow in a pipe is 0·4 over the greater part of the cross section. On entering the transition region near the walls K_0 falls to a value near 0·2. There is a considerable difference between the flow in a pipe and the type of flow with which we are concerned and in view of their stirring mechanism it seems not unreasonable to assume $K_0 = 0{\cdot}4$ for the experimental conditions of Langstroth and Gillespie. A value of this order does account for the magnitude of the deposition in their experiments.

We also take $d\bar{V}/dx$, the velocity gradient of the mean motion, to be constant within the layer, that is, we assume a linear velocity distribution. This is not correct, but the variation to be expected is small compared with the dependence of D_e on x^2, which arises because the mixing length is proportional to the distance from the wall.

Superposed on eddy diffusion we have molecular diffusion, with coefficient D. The combined effect of eddy diffusion and molecular diffusion gives a diffusion coefficient equal to $D_e + D$. In a gravitational field the diffusion process is affected by sedimentation but this can easily be taken into account.

It is obvious that in considering the rate of deposition over each side we can neglect the effect of the other sides. Over the vertical sides of the cube the rate of deposition is governed solely by the processes of eddy and molecular diffusion. Taking axis Ox perpendicular to the side and into the aerosol, x being zero at the side, the concentration within the layer $0 \leqslant x \leqslant \delta$ is given by

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left[(K_e x^2 + D)\frac{\partial c}{\partial x}\right]. \qquad \ldots\ldots(10)$$

At the top and the base of the cube the concentration within the layer $0 \leqslant x \leqslant \delta$ is given by

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left[(K_e x^2 + D)\frac{\partial c}{\partial x}\right] - V\frac{\partial c}{\partial x} \qquad \ldots\ldots(11)$$

and

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\left[(K_e x^2 + D)\frac{\partial c}{\partial x}\right] + V\frac{\partial c}{\partial x} \qquad \ldots\ldots(12)$$

respectively, the last term in each equation being due to the effect of sedimentation, V being the velocity of fall of the particles. Conditions at the paddle face are rather more difficult to assess, but as its area is only 10% of the total we neglect its contribution.

We now solve equation (10), by assuming a solution of the form $c = X_1(x)T_1(t)$, leading to

$$\frac{1}{T_1}\frac{dT_1}{dt} = -\alpha \quad \ldots\ldots(13), \qquad \frac{1}{X_1}\frac{d}{dx}\left[\{K_e x^2 + D\}\frac{dX_1}{dx}\right] = -\alpha, \quad \ldots\ldots(14)$$

where α is a constant determined by the conditions at the wall.

From equation (13) it follows that $\partial c/\partial t = -\alpha c$ and in particular $d\bar{c}/dt = -\alpha\bar{c}$. In equation (14) write $x = (D/K_e)^{1/2}z$, giving

$$(z^2+1)\frac{d^2X_1}{dz^2} + 2z\frac{dX_1}{dz} + \frac{\alpha}{K_e}X_1 = 0.$$

We neglect $\alpha X_1/K_e$, anticipating that it will be small compared with the other terms. Now $X_1 = 0$ at $x = 0$ and to fix the scale of X we make $X_1 = 1$ at $x = \delta$, so that $X_1 = \tan^{-1}[x(K_e/D)^{1/2}]/\tan^{-1}[\delta(K_e/D)^{1/2}]$. We shall see later that in the experiments of Langstroth and Gillespie $\delta(K_e/D)^{1/2}$ was so large that the denominator can be replaced by $\pi/2$.

Using a similar method to solve equations (11) and (12) we obtain $c = X_2(x)T_2(t)$ and

$$[K_e x^2 + D]dX_2/dx - VX_2 = p, \text{ a constant.} \qquad \ldots\ldots(15)$$

In particular $D(dX_2/dx)_0 = p$ and $p\bar{c}/X(\delta)$ is the rate of deposition per unit area on the top side of the cube.

Solving (15) and using $X_2(x) = 1$ at $x = \delta$ and $X_2(x) = 0$ at $x = 0$, we get

$$\frac{1}{V}\log\left(1 + \frac{V}{p}\right) = \frac{1}{(DK_e)^{1/2}}\frac{\pi}{2}. \qquad \ldots\ldots(16)$$

Similarly, for (12), $c = X_3(x)\,T_3(t)$, giving

$$[K_e x^2 + D]dX_3/dx + VX_3 = q, \text{ a constant.} \qquad \ldots\ldots(17)$$

In particular $D(dX_3/dx)_0 = q$ and $q\bar{c}/X_3(\delta)$ is the rate of deposition per unit area on the base of the cube.

Solving (17) and using $X_3(x) = 1$ at $x = \delta$ and $X_3(x) = 0$ at $x = 0$, we get

$$-\frac{1}{V}\log\left(1 - \frac{V}{q}\right) = \frac{1}{(DK_e)^{1/2}}\frac{\pi}{2}. \qquad \ldots\ldots(18)$$

From the rate of loss to the sides we get

$$\frac{d\bar{c}}{dt} = -\frac{D}{L}\left[\frac{4}{X_1(\delta)}\left(\frac{dX_1}{dx}\right)_0 + \frac{1}{X_2(\delta)}\left(\frac{dX_2}{dx}\right)_0 + \frac{1}{X_3(\delta)}\left(\frac{dX_3}{dx}\right)_0\right]\bar{c}$$

or

$$\alpha = \frac{1}{L}\left[\frac{4(DK_e)^{1/2}}{\pi/2} + p + q\right] \qquad \ldots\ldots(19)$$

where p and q are given by equations (16) and (18), from which

$$p+q=V\coth\{V\pi/4(DK_e)^{1/2}\}.$$

Equation (19) can then be written

$$\alpha=\frac{1}{L}\left[\frac{8(DK_e)^{1/2}}{\pi}+V\coth\left(\frac{V\pi}{4(DK_e)^{1/2}}\right)\right]. \qquad \text{......(20)}$$

The velocity V is given by equation (9).

For particles of radius $0\cdot5\,\mu$, as in Langstroth and Gillespie's experiments with turbulent aerosols, $D=3\times10^{-7}\,\text{cm}^2\,\text{sec}^{-1}$. It now remains to determine K_e, which requires the velocity gradient of the mean flow near the sides. We idealize the flow, as far as the neighbourhood of each side is concerned, to turbulent flow over a flat plate in an infinite fluid moving parallel to the plate with mean velocity $\bar{u}$. The total drag on such a plate is given by (cf. Dodge and Thompson 1937, p. 324) $F=0\cdot073\Delta(\rho\bar{u}^2/2)(\eta/\rho\bar{u}l)^{1/5}$ where ρ is the density of the medium, Δ the area of the plate and l the length of the plate in the direction of mean motion. In general this result holds only when the Reynolds' number is less than 10^5, because the derivation is based on the Blasius resistance formula for smooth pipe flow. However, when experimental conditions ensure turbulence practically right from the leading edge of the flat plate the formula remains in good agreement with experiment up to a Reynolds' number of 3×10^6. The conditions in the experiments of Langstroth and Gillespie enable us to assume turbulence all along the sides, and moreover in our idealized flow the maximum value of the Reynolds' number is only of the order 10^5, for mean velocities of flow up to 50 m/min. Also $F=\eta\Delta(\partial\bar{u}/\partial x)_{x=0}$, giving

$$\left(\frac{\partial\bar{u}}{\partial x}\right)_{x=0}=\frac{0\cdot073}{\eta}\frac{\rho\bar{u}^2}{2}\left(\frac{\eta}{\rho\bar{u}l}\right)^{1/5} \qquad \text{......(21)}$$

which for the Langstroth and Gillespie experiments, with $\bar{u}$ in cm. sec^{-1}, leads to $(\partial\bar{u}/\partial x)_{x=0}=7\cdot2\times10^{-2}\times(\bar{u})^{9/5}\,\text{sec}^{-1}$. This implies that δ must be of order $14\,\bar{u}^{-0\cdot8}$ cm.

We now evaluate α for the maximum $\bar{u}$, namely 50 m/min., and mean particle radius $0\cdot5\,\mu$. The thickness δ is then at its smallest, and is of order $\frac{1}{4}$ cm. From (21), $(\partial\bar{u}/\partial x)_{x=0}=2\cdot1\times10^2\,\text{sec}^{-1}$, so that $K_e=33\,\text{sec}^{-1}$; also $V=3\cdot8\times10^{-3}$ cm. sec^{-1} and $\alpha=7\cdot1\times10^{-3}\,\text{min}^{-1}$.

From these it follows that our assumption $\alpha X_1/K_e\ll2z\,dX_1/dz$ is satisfied, near the wall, by a factor of order 10^5, while it is still true at the edge of the layer, though here only by a factor of 10. Our assumption that $\delta(K_e/D)^{1/2}$ is large is also verified, since this quantity is certainly greater than 10^3.

For particle radius of $0\cdot5\,\mu$, and $\bar{u}$ measured in cm. sec^{-1}, we can write

$$\alpha=8\cdot04\times10^{-5}(\bar{u})^{9/10}+2\cdot04\times10^{-3}\coth\,[51/(\bar{u})^{9/10}]\ \text{min}^{-1}$$

which behaves almost linearly as a function of $\bar{u}$ over the range of values of $\bar{u}$, with which we are concerned, except for $\bar{u}<5$ cm. sec^{-1}, and can be replaced by

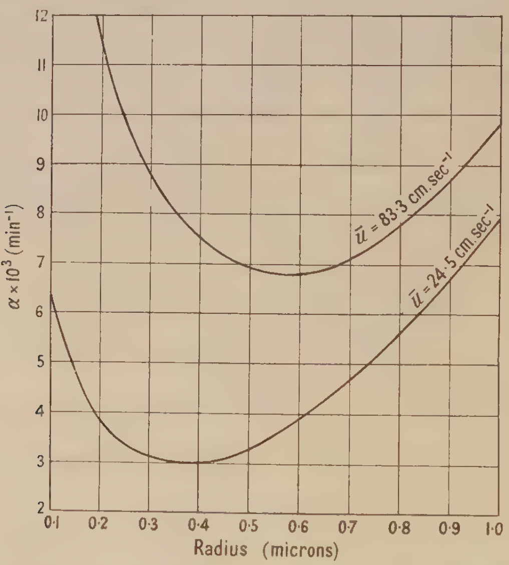
$$\alpha=(2\cdot04+0\cdot061\,\bar{u})10^{-3}\,\text{min}^{-1}\ \text{for}\ \bar{u}>5\ \text{cm. sec}^{-1}. \qquad \text{......(22)}$$

In general, if $\bar{u}$ is measured in cm. sec^{-1} and r in microns then

$$\alpha=5\cdot0\times10^{-5}\left(\frac{1+0\cdot09/r}{r}\right)^{1/2}(\bar{u})^{9/10}$$
$$+6\cdot4\times10^{-3}r^2\left(1+\frac{0\cdot09}{r}\right)\coth\left[\frac{2\cdot57\times10^2r^{5/2}(1+0\cdot09/r)^{1/2}}{(\bar{u})^{9/10}}\right]\text{min}^{-1}. \qquad \text{......(23)}$$

At $\bar{u}=0$ the slope $\partial\alpha/\partial\bar{u}$ is infinite, which is consistent with the experimental results of Langstroth and Gillespie. To obtain agreement with their results for larger $\bar{u}$ we must consider the variation of α with particle size, for it is known from their observations that the mean particle radius decreases with increased turbulence, and this has not been considered in the derivation of equation (22).

If r is constant, then for $\bar{u}$ greater than 5 cm. sec^{-1} the increase of α with $\bar{u}$ is approximately linear. We now consider the variation of α with r for constant $\bar{u}$; typical graphs are shown in the Figure. For a given $\bar{u}$, α has a minimum value for $r=r_{\min}(\bar{u})$. For $\bar{u}$ greater than 5 cm. sec^{-1}, but near the lower end of the range, the mean radius $\bar{r}$ of the particles is greater than $r_{\min}(\bar{u})$, and as $\bar{u}$ increases $\bar{r}$ decreases, thus causing a decrease in the average α. The overall variation of α as $\bar{u}$ increases thus consists of a practically linear increase with $\bar{u}$ as a direct result of the turbulence, together with a decrease depending on the reduction of mean particle radius as an indirect result of the turbulence. The magnitude of this decrease cannot be found without a study of the size distribution, but we can say that over the lower range of $\bar{u}$ the net increase of α with $\bar{u}$ is less than linear. As $\bar{u}$ increases we get a range of values of $\bar{u}$ for which $\bar{r}\simeq r_{\min}(\bar{u})$, and over this range α will increase linearly with $\bar{u}$. Towards the maximum end of the range of $\bar{u}$, $\bar{r}<r_{\min}(\bar{u})$ and the overall increase of α with $\bar{u}$ is greater than linear. The general variation of α with $\bar{u}$ deduced from theory is therefore in agreement with the variation as obtained experimentally by Langstroth and Gillespie.



Variation of α with particle radius for different rates of stirring, calculated from equation (23).

The variation of the deposition rate with size, demonstrated in the Figure, is due to the influence of diffusion, which deposits the smaller particles most readily, and ordinary sedimentation, which is most rapid for large particles. The minimum of the deposition, in the neighbourhood of $\frac{1}{2}$ micron radius for a velocity of air motion of 50 m/min., explains why the stirred aerosols took up a size distribution with a peak in this neighbourhood (cf. Figure 3, Langstroth and Gillespie 1947), whereas in a still aerosol the most common radius soon becomes much larger, roughly one micron.

§ 5. EQUATIONS FOR THE DECAY OF MASS AND PARTICLE NUMBER

Langstroth and Gillespie treated the decay of the mass concentration m and particle concentration n by the equations

$$\frac{dm}{dt} = -\alpha m; \quad \frac{dn}{dt} = -\beta n - Kn^2$$

with the solution

$$m = m_0 e^{-\alpha t}; \quad \frac{1}{n} = \frac{e^{\beta t}}{n_0}\left(1 + \frac{Kn_0}{\beta}(1 - e^{-\beta t})\right) \qquad \ldots\ldots (24)$$

over any interval in which α, β and K can be assumed constant.

From our interpretation of the various terms in these equations we can suggest the dependence of α, β and K on particle size and can relate this to m/n, *provided that the size distribution does not alter in shape.* It is obvious, however, from Figure 3 of Langstroth and Gillespie (1947), that the shape of the size–frequency curve of the particles alters at much the same rate as the average size, at any rate under their conditions. Thus it is doubtful whether much is to be gained by using merely the dependence on mean radius. To introduce the size distribution would lead to complicated equations and the gain in accuracy would be small, as is clear from the success of the Langstroth–Gillespie equations. We have confined our attention to showing that the dependence on mean radius does not require large changes in the constants.

The coagulation terms are $dn/dt = -K_0 n^2 - K_1 n^2$ where K_0, the rate in still air, is almost independent of particle size, though it does depend a little on the shape of the size–frequency curve. K_1 is the coagulation rate due to stirring, and is proportional to particle volume, that is, to m/n. The rate of loss by sedimentation is $dm/dt = -\alpha m$, $dn/dt = -\beta n$, with α and β proportional to $(m/n)^{2/3}$. Similar equations hold for the loss to the walls when the air is stirred; the constants are here proportional to $(m/n)^{-1/6}$, the variation of which may be neglected. Hence

$$\frac{dm}{dt} = -\alpha\left(\frac{n_0}{m_0}\right)^{2/3}\left(\frac{m}{n}\right)^{2/3} m - \alpha' m,$$

$$\frac{dn}{dt} = -\beta\left(\frac{n_0}{m_0}\right)^{2/3}\left(\frac{m}{n}\right)^{2/3} n - \beta' n - K_0 n^2 - K_2 \frac{n_0}{m_0} mn.$$

These equations do not appear to have a simple solution. They could, of course, be solved numerically for any given set of the parameters. Fitting to the experimental data would therefore be laborious, especially in view of the six arbitrary parameters. We have confined our attention to the still-air case, for which α', β' and K_2 are certainly zero. Although an exact solution for the simplified equations is still not known, it is possible to get an approximation, sufficient for the analysis of the Langstroth–Gillespie results, by expanding the correction terms as a power series in the time. We find

$$\ln(m/m_0) = -\alpha t - \frac{\alpha\beta}{3}\left(1 - \frac{\alpha}{\beta} + \frac{K_0 n_0}{\beta}\right) t^2 + O(t^3)$$

and

$$\frac{n_0}{n} + \frac{K_0 n_0}{\beta}\left[1 + \frac{\beta^2}{3}\left(1 - \frac{\alpha}{\beta} + \frac{K_0 n_0}{\beta}\right) t^2 + O(t^3)\right]$$

$$= \left(1 + \frac{K_0 n_0}{\beta}\right) \exp\left[\beta t + \frac{\beta^2}{3}\left(1 - \frac{\alpha}{\beta} + \frac{K_0 n_0}{\beta}\right) t^2 + O(t^3)\right].$$

These contain no more arbitrary parameters than (24). They give just as good a fit to the still-air results as was obtained by Langstroth and Gillespie with equation (24); α is slightly bigger and β slightly less, but the changes are less than 20%. The fit to experiment is certainly not obviously better than was obtained by Langstroth and Gillespie, and the reason is probably the change in the particle size distribution as sedimentation and coagulation proceeds.

ACKNOWLEDGMENTS

Acknowledgment is made to the Chief Scientist, Ministry of Supply, and to the Controller, H.M. Stationery Office, for permission to publish this paper.

REFERENCES

Dodge, R. A., and Thompson, M. J., 1937, *Fluid Mechanics* (New York : McGraw-Hill).
Fuchs, N., 1934, *Phys. Z. Sowjet*, **171**, 199.
Langstroth, G. O., and Gillespie, T., 1947, *Canad. J. Res.*, B **25**, 455.
Langstroth, G. O., Gillespie, T., and Pearce, R. M., 1949, *Chem. Rev.*, **44**, 337.
Millikan, R. A., 1923, *Phys. Rev.*, **22**, 1.
Patterson, H. S., Whytlaw-Gray, R., and Cawood, R., 1927, *Proc. Roy. Soc.* A, **116**, 540.
Smoluchowski, M. V., 1917, *Z. phys. Chem.*, **92**, 129.
Taylor, G. I., 1935 a, *Proc. Roy. Soc.* A, **151**, 421; 1935 b, *Ibid.*, **151**, 444.
Tikhomirov, M. V., Tunitzki, N. N., and Petrjanov, J. B., 1942, *Acta Physicochim.*, **17**, 185.
Tunitzki, N. N., 1946, *J. Phys. Chem. U.S.S.R.*, **20**, 1137.

On Oscillations in Electron Streams

By R. Q. TWISS
Services Electronics Research Laboratory, Baldock, Herts.

MS. received 11*th January* 1951, *and in amended form* 2nd *March* 1951

ABSTRACT. A critical discussion is given of the conventional procedure for determining whether a particular electron stream is unstable. It is shown that in some cases where this analysis suggests instability the stream is stable, while in others, where instability does occur, neither the frequency nor the rate of build up of oscillations are as predicted. The physical causes of instability in an electron stream are also considered.

§1. INTRODUCTION

The existence of plasma oscillations was first established by Tonks and Langmuir (1929) in the course of their investigations into discharge tube phenomena. On the assumption that thermal motion was slow enough to be neglected, Tonks and Langmuir showed that an electron–ion plasma would oscillate, when disturbed, about its electrically neutral state at an angular frequency, in M.K.S. units, of $\omega_0=(e^2N_0/\epsilon_0 m)^{1/2}$, where N_0 is the number of particles of charge $-e$ and mass m in unit volume.

The development of a similar theory for oscillations in electron streams has been undertaken by different writers (Bailey 1948, Bohm and Gross 1949, Malmfors 1950) from several points of view; but most of these treatments have been based upon a substitution analysis in which one looks for solutions of the general form $\exp i(\mathbf{k}.\mathbf{x}+\omega t)$ and obtains an equation relating k and ω, the so-called dispersion or eigen-equation. The roots of this equation viewed as an equation for ω for fixed $\mathbf{k}$ are interpreted as giving the oscillation frequencies of the plasma in a particular spatial mode. Should any of these roots be complex it is concluded that the plasma is unstable and that oscillations will build up from an initial disturbance to a final amplitude limited only by the available non-thermal energy of the plasma.

It is the purpose of this paper to emphasize that this interpretation of the dispersion equation is valid only when the disturbance in the electron stream can be represented by a system of standing waves, a case that rarely arises in moving media. In some cases where this interpretation predicts instability the flow is stable, while in others, where instability does occur, neither the frequency nor the rate of build up of oscillation are given correctly by this procedure.

The error has arisen largely because the substitution analysis does not specifically include the boundary and initial conditions, without which the dispersion equation is simply a mathematical expression of little physical meaning.

Often it is not instability but amplification that is taking place, a difference that is of considerable importance when we come to consider the actual phenomena that can take place in a discharge tube or solar atmosphere. The distinction between these two processes can be made clear as follows. Suppose that at time $t=0$ we impose a disturbance on a restricted region of the plasma, which is so small that non-linear effects do not appear in any finite time–space interval. If after an arbitrarily long time the disturbance in a neighbouring region is finite in amplitude but increases exponentially as $\exp \boldsymbol{\alpha}.\mathbf{x}$, where $\boldsymbol{\alpha}$ has a positive real part, as we move away from the initially disturbed region, then we say that the medium is amplifying. However if the disturbance at a fixed distance from the point of amplification increases exponentially as $\exp \alpha t$, then we say that the medium is unstable.

A full discussion of the conditions for instability in a plasma has been given in two recent papers by Bohm and Gross (1949) hereafter referred to as B–G; to shorten our discussion we shall employ a similar notation and freely use their results. We shall first give a brief outline of the use of the substitution analysis and criticize its conventional interpretation. This qualitative analysis will then be illustrated by a specific case which is in fact stable although the procedure that we are criticizing would suggest instability. Finally we shall briefly examine the conclusions that can be drawn from this discussion as to the actual conditions under which electron streams might exhibit instability.

§ 2. THE SUBSTITUTION ANALYSIS

In theoretical treatments of the electron plasma it is quite usual to confine attention to the so-called longitudinal plasma oscillations. This means, in effect, that one ignores the magnetic component of the electromagnetic field and assumes that the system is described completely by the Poisson equation

$$\nabla.\mathbf{D}=\sum_{s=1}^{N}\rho_s \qquad \ldots\ldots(1)$$

the set of continuity equations

$$\frac{\partial\rho_s}{\partial t}+\nabla(\rho_s.\mathbf{v}_s)=0 \quad (s=1\ldots N) \qquad \ldots\ldots(2)$$

and the Lorentz force equations

$$\left(\frac{\partial}{\partial t}+\mathbf{v}_s.\nabla\right)m\mathbf{v}_s=-e\mathbf{E}\quad(s=1\ldots N),\qquad\ldots\ldots(3)$$

where, following Bohm and Gross, we have ignored the effect of the short-range Coulomb collisions.

In these equations, which describe the propagation of longitudinal waves in an electron stream composed of N separate velocity beams, ρ_s is the space-charge density of the sth beam, $\mathbf{v}_s$ is the velocity of the electrons of the sth beam, $-e$ is the electronic charge, and the summation in equation (1) is taken over the N beams constituting the stream.

This division into transverse and longitudinal waves is assumed to be approximately valid as long as the electron velocities are all very small compared with the velocity of light, and it is rigorously justifiable even for relativistic velocities in the unidimensional case.

On the assumption that all the disturbances are small, the linear approximation to equations (1) to (3) may be used which yields

$$\left.\begin{aligned}&\nabla.\mathbf{D}=\sum_{s=1}^{N}\rho_s,\quad\frac{\partial\rho_s}{\partial t}+\rho_{0s}\nabla.\mathbf{v}_s+\mathbf{u}_{0s}\nabla\rho_s=0,\\&\left(\frac{\partial}{\partial t}+\mathbf{u}_{0s}.\nabla\right)m\mathbf{v}_s=-e\mathbf{E},\end{aligned}\right\}\qquad\ldots\ldots(4)$$

where the subscript zero identifies the time independent terms in the velocity and charge density of the N electron beams.

One then looks for solutions in which the time and space dependent components of the field and space-charge variables can be represented by a summation over a set of partial waves of the form $\exp i(\mathbf{k}.\mathbf{x}+\omega t)$ and obtains a dispersion equation between k and ω of the form

$$1-\sum_{s=1}^{N}\frac{\omega_{0s}^{2}}{(\omega+\mathbf{u}_{0s}.\mathbf{k})^{2}}=0\qquad\ldots\ldots(5)$$

where ω_{0s} is the angular plasma frequency defined by

$$\omega_{0s}=(-e\rho_{0s}/\epsilon_0 m)^{1/2}.\qquad\ldots\ldots(6)$$

When the velocity distribution in the plasma is continuous, the dispersion equation may be written

$$1-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{d\omega_{0s}^{2}(\mathbf{u}_{0s})}{(\omega+\mathbf{u}_{0s}.\mathbf{k})^{2}}=0,\qquad\ldots\ldots(7)$$

where $d\omega_{0s}^{2}(\mathbf{u}_{0s})=(-e\rho(\mathbf{u}_{0s})/\epsilon_0 m)\,d\mathbf{u}_{0s}$.

Now the dispersion equation (5) is the three-dimensional form of the characteristic equation obtained by Haeff (1949) and independently by Pierce (1948) in their theories of space-charge wave amplification. These writers regard the dispersion equation as an equation for $\mathbf{k}$ in terms of ω; should any of the roots be complex they conclude that amplification is taking place. This interpretation is valid when the electron streams are injected into the interaction space across a material boundary with an initial sinusoidal modulation of angular frequency ω caused either by the shot noise in the electron beams or imposed from an external source.

On the other hand B–G are concerned with the transient response to a particular initial spatial distribution of velocity and charge density modulation, set up by scattering or by ionizing radiation. However, despite these differences in the initial and boundary conditions, the physical mechanism of interaction between the constituent beams of the stream is the same in both cases, which renders it doubtful whether both interpretations of the dispersion equation can lead to physically observable consequences.

The next step in the conventional procedure is to analyse the disturbance into the appropriate normal modes. If the plasma be contained in a large rectangular box, of dimensions (L, M, N) then a general spatial disturbance can be analysed as a triple Fourier series

$$\sum_{l,m,n=-\infty}^{\infty} A_{l,m,n} \exp i\mathbf{k}_{l,m,n}\mathbf{x},$$

where the allowable values of **k** are given by $k_x = 2\pi l/L$, $k_y = 2\pi m/M$, $k_z = 2\pi n/N$.

When the box has infinite dimensions, as assumed by B–G, this triple sum must be replaced by the appropriate triple integral.

The allowable values of **k** thus range from a minimum fixed by the inverse dimensions of the plasma region to a maximum that would be infinite were not the plasma composed of discrete particles, and which is anyway assumed to be so high as easily to cover all the cases of interest.

For given **k** the dispersion equations (5) and (7) yield a set of possible values for ω some of which may be conjugate complex and of the form $\omega = \Omega_0 \pm i\alpha$ say. The wave associated with the negative choice of sign is of the form $A_{l,m,n} \exp \alpha t \exp i(\mathbf{k}.\mathbf{x} + \Omega_0 t)$ which grows exponentially with the time, and it is concluded that the stream, in this particular mode, is unstable.

Clearly, however, this conclusion is valid only when the disturbance can be represented by a system of standing waves, for it is based upon the tacit assumption that the expansion into the set of normal spatial modes is time invariant. Accordingly it is inapplicable to an electron stream, for the longitudinal plasma oscillations, being localized in the stream, are convected away by the moving charge.

It is true that, for the infinite plasma which is analysed by B–G, this procedure gives a formally correct result, if we assume that the initial disturbance, which in the simplest case might be of the form $\rho_0 \exp ik_0 z$, is defined in the range $-\infty < z < \infty$. However, in this case it is impossible to distinguish on purely mathematical grounds between amplification and instability. If the medium is unstable then the initial disturbance at a point z will build up exponentially with time, the general spatial form of the initial disturbance being preserved. But this is just what would be observed if the initial disturbance were amplified by interaction between the various streams of the plasma as it was convected along by the moving charge.

The ambiguity can immediately be resolved however if we assume, as is always the case in a physically realizable situation, that the initial disturbance is confined to a finite region bounded, let us say, by planes at $z = 0$ and $z = d$. There is no need to assume that these boundaries are material ones, it is sufficient that the electron streams entering the interaction space across these boundaries be modulated in a manner independent of the initial disturbance in the region $0 < z < d$.

Up to now of course we have only cast doubt on the validity of a particular mathematical procedure, we have not shown that the conclusions derived by it may be wrong. To prove this we will now consider quantitatively a specific case, which will serve the additional purpose of throwing light on the general physical conditions that can lead to instability in electron streams.

§3. THE TWO-VELOCITY ELECTRON STREAM

We consider the case where the stream consists of two beams of charged particles moving with well-defined velocities u_{01}, u_{02}, where $u_{02} > u_{01}$, and densities ρ_{01}, ρ_{02} respectively. This has been analysed by B–G when the two streams are moving in opposite directions; they point out that, on physical grounds, we should expect this flow to be unstable since the reverse beam provides a feedback mechanism. However, the straightforward application of their procedure implies that the flow should be unstable even when the beams are moving in the same direction, as we shall now show.

From equation (5) the dispersion equation for the unidimensional two-beam case may be written

$$1-\left[\frac{\omega_{01}{}^2}{(\omega+u_{01}k)^2}+\frac{\omega_{02}{}^2}{(\omega+u_{02}k)^2}\right]=0. \qquad \ldots\ldots(8)$$

When $\omega_{01}{}^2=\omega_{02}{}^2=\omega_0{}^2$ this equation reduces to a biquadratic in ω with the solution

$$\omega=-\left(\frac{u_{01}+u_{02}}{2}\right)k\pm\omega_0\left[1+\left(\frac{k(u_{02}-u_{01})}{2\omega_0}\right)^2\pm\left\{1+4\left(\frac{k(u_{02}-u_{01})}{2\omega_0}\right)^2\right\}^{1/2}\right]^{1/2} \qquad \ldots\ldots(9)$$

two of these roots being conjugate complex if $0<k<2\sqrt{2}\omega_0/(u_{02}-u_{01})$.

As we pointed out above, the minimum value of k is fixed by the inverse dimensions of the plasma so that it would appear that the flow should be unstable if its length satisfies the inequality

$$L>\frac{2\pi}{2\sqrt{2}\omega_0/(u_{02}-u_{01})}. \qquad \ldots\ldots(10)$$

A similar result holds when $\omega_{01}{}^2\neq\omega_{02}{}^2$, although the critical value of L will be different.

For the experimental tube built by Haeff, where the electron beams had energies of 200 and 250 volts respectively, and where the plasma frequency was about 2×10^8 c/s., this critical length is found to be about 1 millimetre, in contrast to the fact that stable tubes of lengths greater than 50 cm. have been built with appreciable amplification.*

To justify our contention that the unidimensional flow is stable, provided that $u_{01}>0$, we will obtain the transient solution under arbitrary initial conditions. Since we are not interested in the exact nature of the solution and since the mathematical procedure is a familiar one only an outline is given in the text, the more important details being given in an Appendix.

---

* It is true that such devices are unstable if their cross section is made too large but, as we argue below, the cause of instability in this case is the presence of a backward travelling transverse electromagnetic wave produced by reflections from the far end of the amplifying tube. In the infinite unidimensional case this wave is not excited by the longitudinal oscillations.

The most direct method of finding the transient response of a linear system to an arbitrary initial signal is to use the Laplace transform analysis as this incorporates the initial conditions explicitly in the solution.

When analysing the unbounded plasma our present purposes would be served by assuming that there is an arbitrary initial spatial distribution of charge density and velocity modulation of both beams, $\rho_1(z,0)$, $\rho_2(z,0)$, $v_1(z,0)$, $v_2(z,0)$, in the region $0<z<d$, and that outside this range the disturbance is initially zero. However the finite plasma into which the electron beams are injected from a boundary is of at least equal interest and to cover this case as well we may imagine that the beams possess a time dependent modulation at the initial surface $\rho_1(0,t)$, $\rho_2(0,t)$, $v_1(0,t)$, $v_2(0,t)$.

When the beams are moving in the same direction and are injected from the same boundary it is plausible that these time dependent modulations are independent of each other and of the initial spatial distribution of modulation. When the beams are moving in opposite directions, however, this assumption is clearly invalid and a different procedure is necessary. From this point on, the solution is carried out in the Appendix. From the results derived there, it is easy to see the general nature of the propagation.

To begin with let us consider the case of the infinite plasma where the initial disturbance is confined to the region $0<z<d$. Then we see that the plasma oscillations are indeed localized in the plasma being convected about by the moving charge. This follows from the fact that the disturbance at a point $z>0$ is zero after the time $t \geqslant z\ u_{01}$ taken by the electrons of the slower beam which cross the surface $z=0$ at $t=0$ to reach z, while the disturbance at a point $z>d$ is zero until a time $t \geqslant (z-d)\ u_{02}$ when the faster moving electrons, which left $z=d$ at $t=0$, arrive.

Hence as time goes on, the disturbance originally located in the region $0<z<d$ is convected away until at time t it is localized in the region $u_{01}t<z<d-u_{02}t$. Furthermore, because of the interaction between the two electron beams, this disturbance is continually being amplified at a rate that depends on the space-charge density and, to some extent, on the exact nature of the initial conditions.

At any point within a finite distance of $z=0$, the disturbance is always finite, while after a time $t>z\ u_{01}$ it is identically zero. Hence in the sense used in this paper the flow is stable. If an observer were to move with the disturbance at a speed u, where $u_{02}>u>u_{01}$, then he would see a disturbance that increased indefinitely with time and would conclude that the stream was unstable. However, in this case the two beams would be flowing in opposite directions with respect to the observer, a condition that on physical grounds we would expect to lead to instability. In the infinite uniform stream, where one frame of reference is on the same basis as any other, we see that the ambiguity as to whether amplification or instability is taking place is still unresolved; but this is a purely academic difficulty. In any case that could arise in the laboratory there is always a preferred frame of reference: that in which the plasma boundaries are at rest. In a stellar atmosphere it is possible for the surfaces of discontinuity* to be moving with respect to each other, but in no case would all frames of reference be on an equal footing.

It may also be noted that one can only establish an infinite stream which is either inherently unstable or amplifying, by very artificial means, such as suddenly ionizing a flow of neutral gas atoms.

---

* These would be the boundaries of those regions in which there was a rapid change of plasma density, or magnetic field, or electron temperature.

Incidentally the analysis given in the Appendix demonstrates the gain in physical insight into the nature of the solution that is obtained by using the single-sided Laplace transform analysis rather than the conventional double sided Fourier transform. Although the solution in the Appendix has been carried through for a single Laplace component of the initial disturbance the fact that at point z the disturbance is identically zero for all $z>d$ is immediately apparent. If we had used the Fourier analysis however, this result could only be made apparent by summing over all the separate Fourier components: the solution for a single component is liable to lead to the false interpretation that the medium is unstable.

The nature of the solution for the bounded stream is very similar to that for the infinite case when the two beams are moving in the same direction. Thus the disturbance at point $z<d$ is quite independent of d or of the nature of the boundary there. If a time dependent modulation be applied at $t=0$ to the electron beams entering the plasma then it is shown in the Appendix that the disturbance for $t>z/u_{01}$ is just that predicted by the steady state theory of Haeff and Pierce.

When the electron beams are moving in opposite directions instability is likely to result, as discussed in the Appendix. However, the expressions for the associated oscillation frequencies and rates of build up are, in general, very complex. Only in very special circumstances would they coincide with those predicted by the procedure appropriate for the standing-wave case.

§4. INSTABILITY IN A THREE-DIMENSIONAL PLASMA

From physical intuition one would expect that instability, in the sense used in this paper, could only arise in a plasma* if a mechanism existed by which energy could continually be fed back. For the longitudinal plasma oscillations it appears, by a simple generalization of the results proved above for the two-beam electron stream, that such a mechanism can only be provided by an electron beam moving in the reverse direction to the main flow. In a laboratory device such beams might be produced by back scattering from an anode, by injection from a separate cathode or, as in a reflex klystron, by reflection from a potential barrier.†

In the general case where all directions of propagation of the field and space-charge waves are permitted, the distinction between longitudinal and transverse waves is less easy to draw, since in general any wave will possess all components of both E and H and the analysis becomes appreciably more complex. For one thing an additional cause of instability now appears because, at a surface of discontinuity or boundary of the plasma, a backward travelling field wave ‡ will be reflected on the incidence of a forward travelling field or space-charge wave. This is particularly likely to lead to instability when the incident space-charge wave is amplified with a forward gain greater than the backward attenuation of the field wave, as is well known from the experimental behaviour of travelling wave and space-charge wave amplifiers.

It may be thought that the interaction of these waves with the electron stream is likely to be negligible, except in special cases, because they propagate at a speed so much greater than the electron velocity. However, this is not so in electron tubes which are well below cut-off to the field waves, since the axial component of the electric vector can here be large.

---

* In what follows we shall understand a plasma to mean any ionized medium that is electrically neutral, whether composed of well-defined beams or in thermal equilibrium.

† The existence of such a reverse flow is a necessary, not a sufficient, condition for the growth of such oscillations, cf. the case of a plasma in thermal equilibrium.

‡ By field waves we mean those waves which propagate, at speeds of the order of magnitude of c, and which still exist when the space-charge density tends to zero. The space-charge waves on the other hand propagate at approximately the velocity of one or other of the electron beams and vanish as the space-charge density tends to zero.

It should be noted that a surface of discontinuity will exist in an amplifying plasma, in the neighbourhood of the region where non-linear effects become important and alter the average or time independent nature of the plasma flow. Field waves reflected from this region provide feedback that may lead to instability which in turn can alter the position of the region of non-linearity.

It seems plausible that instability in a plasma can only arise in the presence of a reverse electron beam or of a field wave reflected from a boundary or surface of discontinuity. However, the proof of this lies beyond the scope of the present paper, it is advanced here simply as a hypothesis that is supported by the special cases considered above.

§ 5. CONCLUSIONS

The principal aim of this paper has been to cast doubt upon the value and indeed upon the validity of the commonly used procedure for finding the resonant frequencies of a plasma containing moving charge, where one solves the characteristic or dispersion equation for ω for a fixed value of the propagation vector $\mathbf{k}$. We have argued that only in very special cases can one deduce from the nature of these roots the conditions for stability, or the frequencies and rates of build up of the free oscillations of the plasma which would actually be observed. This conclusion was confirmed for the special case of a two-beam electron stream, which emphasized the necessity of taking into account the initial and boundary conditions if physically meaningful results are to be obtained.

We showed that the criticized procedure is valid only for those phenomena where the disturbance can be expressed as the sum of a system of independent standing waves, e.g. the oscillations of a fixed elastic string or of a stationary plasma. It is likely to be misleading when the disturbance can only be analysed by travelling waves, or when the possible stationary modes are interdependent, as then such an analysis will often predict instability when, in fact, it is amplification that is taking place, a difference that is of considerable physical importance.

For the unidimensional case we showed that the longitudinal plasma oscillation can only become unstable if there is a reverse electron beam to provide a feedback path. In the general case instability can also be caused by a field wave reflected from a surface of discontinuity or a boundary of the plasma, but a more powerful theory is needed before the necessary and sufficient conditions for instability can be established.

ACKNOWLEDGMENT

Acknowledgment is made to the Admiralty for permission to submit this paper for publication.

REFERENCES

Bailey, V. A., 1948, *J. Roy. Soc. N.S.W.*, **82**, 107.
Bohm, D., and Gross, E. P., 1949, *Phys. Rev.*, **75**, 1851.
Haeff, A. V., 1949, *Proc. Inst. Radio Engrs., N.Y.*, **37**, 4.
Malmfors, K. G., 1950, *Ark. Fys.*, I, **30**, 569.
Pierce, J. R., 1948, *Bell Syst. Tech. J.*, **28**, 33.
Tonks, L., and Langmuir, I., 1929, *Phys. Rev.*, **33**, 195.

APPENDIX

THE TRANSIENT RESPONSE OF A TWO-VELOCITY ELECTRON STREAM

In this Appendix we shall obtain the solution for the two-velocity electron stream considered in the text with the aid of two successive Laplace transformations defined by the relations

$$L_t\{f(z,t)\}=f^t(z,\omega)=\int_0^\infty f(z,t)\exp(-i\omega t)\,dt,$$

$$L_t^{-1}\{f^t(z,\omega)\}=f(z,t)=\frac{1}{2\pi}\int_{-\infty-i\gamma}^{\infty-i\gamma} f^t(z,\omega)\exp i\omega t\,d\omega$$

and
$$L_z\{f^t(z,\omega)\}=f^{t*}(k,\omega)=\int_0^\infty f^t(z,\omega)\exp(-ikz)\,dz,$$

$$L_z^{-1}\{f^{t*}(k,\omega)\}=f^t(z,\omega)=\frac{1}{2\pi}\int_{-\infty-i\gamma}^{+\infty-i\gamma} f^{t*}(k,\omega)\exp ikz\,dk,$$

where all the singularities of $f^{t*}(k,\omega)$ and $f^t(z,\omega)$ lie above the lines $\mathrm{Im}(k)+\gamma=0$ and $\mathrm{Im}(\omega)+\gamma=0$ respectively.

To simplify the algebra we shall not consider the response to a general disturbance but shall assume instead that it is caused by an arbitrary modulation of the density of one of the electron beams which is confined to the region $0<z<d$. That is, we assume that $\rho_2(z,0)=v_2(z,0)=v_1(z,0)=0$ and that $\rho_1(z,0)$ can be represented by a Fourier integral

$$\rho_1(z,0)=\int_{-\infty}^{\infty} K(k_0)\exp ik_0z\,dk_0, \qquad \ldots\ldots(\mathrm{A}\,1)$$

where $\rho_1(z,0)=0$, $-\infty<z<0$ and $d<z<\infty$. We have immediately that the Laplace transform of $\rho_1(z,0)$ with respect to z is given by

$$\rho_1(k,0)=\int_{-\infty}^{\infty} K(k_0)\left[\frac{1-\exp\{-i(k-k_0)d\}}{i(k-k_0)}\right]dk_0 \qquad \ldots\ldots(\mathrm{A}\,2)$$

and without loss of generality we can take $K(k_0)=K\delta(k_0)$ where $\delta(k)$ is the Dirac δ-function, if k_0 is taken as an arbitrary real number.

In order to cover the case of the bounded plasma where the injected beams possess an arbitrary modulation independent of $\rho_1(z,0)$ we may take

$$\rho_1(0,t)=W\exp i\Omega_0 t \quad (t>0), \qquad \ldots\ldots(\mathrm{A}\,3)$$

so that $\rho_1{}^t(0,\omega)=W/i(\omega-\Omega_0)$ where W is to be taken as equal to zero for the unbounded plasma.

If we now take the Laplace transform of the unidimensional form of equation (4) first with respect to t and then with respect to z, and eliminate $D^t(0,\omega)$ with the aid of the Maxwell equation

$$\frac{\partial D(z,t)}{\partial t}+\sum_{s=1}^{2}[\rho_{0s}v_s(z,t)+u_{0s}\rho_s(z,t)]=0, \qquad \ldots\ldots(\mathrm{A}\,4)$$

we obtain a system of linear algebraic equations which may be solved for $D^{t*}(k,\omega)$ to give $D^{t*}(k,\omega)=$

$$\left\{1-\sum_{s=1}^{2}\frac{\omega_{0s}{}^2}{(\omega+u_{0s}k)^2}\right\}^{-1}\left\{-\frac{K[1-\exp\{-i(k-k_0)d\}]}{ik(k-k_0)(\omega+u_{01}k)}+\frac{Wu_{01}}{i\omega(\omega-\Omega_0)(\omega+u_{01}k)}\right\}$$

where $\omega_{0s}{}^2=-e\rho_{0s}/\epsilon_0 m$.

$D^{t*}(k, \omega)$ is a meromorphic function of k with poles at

$$k=0, \qquad \ldots\ldots(\mathrm{A}\,5), \qquad\qquad k-k_0=0, \qquad \ldots\ldots(\mathrm{A}\,6)$$

$$1-\sum_{s=1}^{2}\frac{\omega_{0s}{}^2}{(\omega+u_{0s}k)^2}=0, \qquad \ldots\ldots(\mathrm{A}\,7)$$

which is $O(k^{-1})$ as $k \to \infty$.

In the case that $\omega_{01}{}^2/u_{01}{}^2=\omega_{02}{}^2/u_{02}{}^2=\omega_0{}^2/|u_{01}u_{02}|=\omega_0{}^2/v_0{}^2$, equation (A 7), which is the dispersion or characteristic equation given by the substitution analysis, can be solved for k to give the four roots

$$k_n=-\frac{\omega(u_{01}+u_{02})}{2v_0{}^2}+(-)^n\frac{\omega_0}{v_0}\left\{1+\left[\frac{(u_{02}-u_{01})\omega}{2v_0\omega_0}\right]^2 + \cos\frac{(2n-1)}{4}\pi\left(2\left\{1+4\left[\frac{(u_{02}-u_{01})\omega}{2v_0\omega_0}\right]^2\right\}\right)^{1/2}\right\}^{1/2} \qquad n=1,2,3,4 \quad \ldots(\mathrm{A}\,8)$$

and the inversion from the k into the z-plane by the integral

$$D^t(z,\omega)=\frac{1}{2\pi}\int_{-i\gamma-\infty}^{-i\gamma+\infty}D^{t*}(k,\omega)\exp ikz\,dk$$

can be performed by transforming the line integral of the separate terms of $D^{t*}(k, \omega)$ into contour integrals around the upper or lower infinite half circle, according to the value of z, and using the theory of residues. We get

$$D^t(z,\omega)=K\left\{\frac{1}{\omega k_0\left[1-\sum\limits_{s=1}^{2}\frac{\omega_{0s}{}^2}{\omega^2}\right]}-\frac{\exp ik_0z}{k_0(\omega+u_{0s}k_0)\left[1-\sum\limits_{s=1}^{2}\frac{\omega_{0s}{}^2}{(\omega+u_{0s}k_0)^2}\right]}\right.$$
$$\left.-\sum_{n=1}^{4}\frac{\exp(ik_nz)\left(\frac{\omega}{u_{01}}+k_n\right)\left(\frac{\omega}{u_{02}}+k_n\right)^2}{k_n(k_n-k_0)\prod\limits_{m\neq n=1}^{4}(k_n-k_m)}\right]$$
$$+W\sum_{n=1}^{4}\frac{\exp(ik_nz)u_{01}\left(\frac{\omega}{u_{01}}+k_n\right)\left(\frac{\omega}{u_{02}}+k_n\right)^2}{\omega(\omega-\Omega_0)\prod\limits_{m\neq n=1}^{4}(k_n-k_m)} \qquad \ldots\ldots(\mathrm{A}\,9)$$

for $d>z>0$, and $D^t(z, \omega)=0$ for $z<0$, and a similar expression, containing terms in $\exp ik_0d$ for $D^t(z, \omega)$ in $d<z<\infty$ where the k_n are given by equation (A 8).

The final step in the solution, the inversion from the ω-plane into the t-plane is complicated by the fact that k_n and hence $D^t(z, \omega)$ are multi-valued functions of ω, with branch points where

$$1+4\left[\frac{(u_{02}-u_{01})\omega}{2v_0\omega_0}\right]^2=0, \qquad \ldots\ldots(\mathrm{A}\,10)$$

that is, at $\omega=\pm i\omega_0\, v_0/(u_{02}-u_{01})$, and where

$$1+\left[\frac{(u_{02}-u_{01})}{2v_0}\frac{\omega}{\omega_0}\right]^2\pm\left[1+4\left(\frac{(u_{02}-u_{01})\omega}{2v_0\omega_0}\right)^2\right]^{1/2}=0, \qquad \ldots\ldots(\mathrm{A}\,11)$$

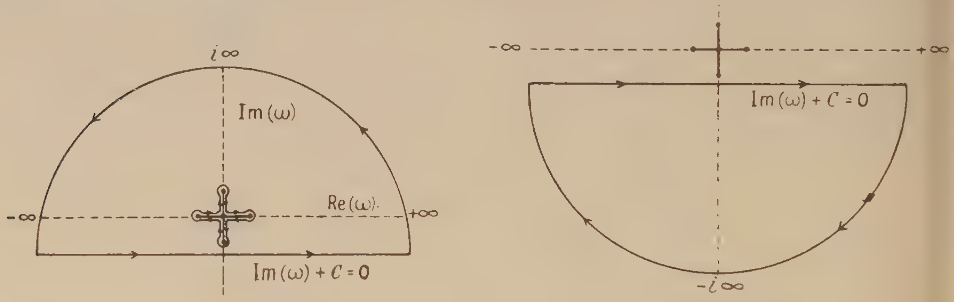
that is, at $\omega=0$, $\omega=\pm\omega_0 v_0\,(u_{02}-u_{01})$, and to make these functions single valued it is necessary to introduce a system of cuts into the complex ω-plane as shown n Figure 1. The contribution to $D(z, t)$ of the various terms of $D^t(z, \omega)$ is then found by transforming the associated line integral into an integral around a closed contour which will be that either of Figure 1 (*a*) or Figure 1 (*b*), the actual choice being determined by the requirement that the contribution to the integral from the appropriate infinite half circle be zero.

When $0<z<d$ the first two terms of $D^t(z, \omega)$ in equation (A 9) are evaluated from the contour of Figure 1 (*a*) for $t>0$ and from that of Figure 1 (*b*) for $t<0$, but the critical value of t is different for the remaining terms as $\exp ik_n z$ is a function of ω. In fact, for sufficiently large ω, $k_1, k_3 \sim -\omega/u_{01}$ and $k_2, k_4 \sim -\omega/u_{02}$, so that the terms in $D^t(z, \omega)$ involving the factors $\exp ik_2 z$, $\exp ik_4 z$ contribute nothing to $D(z, t)$ until $t>z/u_{02}$, while it is not until $t>z/u_{01}$ that all the terms of $D^t(z, \omega)$ contribute to $D(z, t)$.

When $0<t<z/u_{02}$ the solution can easily be effected as the first two terms of $D^t(z, \omega)$ are single valued all over the complex ω-plane and their line integral is equal simply to $2\pi i \times$ the sum of the residues at their poles. The contribution of the first term is independent of z and proportional to $\exp \pm i(\omega_{01}^2 + \omega_{02}^2)^{1/2} t$, that is, an oscillation at the plasma frequency. The residues of the second term of equation (A 9) on the other hand are those of the poles given by the zeros of

$$1 - \sum_{s=1}^{2} \frac{\omega_{0s}^2}{(\omega + u_{0s} k_0)^2} = 0,$$

which are just the waves given by the solution of B–G.



(*a*) Contour for $t>z/u_{01}$. (*b*) Contour for $t<z/u_{02}$.

Figure 1. Contours for the inversion of $D^t(z, \omega)$ from the ω-plane into the t-plane.

In the intermediary range $z/u_{02}<t<z/u_{01}$ the expression for $D(z, t)$ is complicated, and though of some interest will not be discussed further here, but for $t>z/u_{01}$ it is simply the steady state response to the initial time dependent density modulation. This follows because the only non-zero contribution to $D(z, t)$ now comes from the pole at $\omega - \Omega_0 = 0$. The contribution from any other pole such as that at $k_n = k_m$ or $k_n = k_0$ or $k_n = 0$ is identically zero since these poles always occur antisymmetrically in pairs, the residue of $f(k_n) \exp(ik_n z)/(k_n - k_0)$ for example, being equal in magnitude but opposite in sign to the residue of $f(k_0) \exp(ik_0 z)/(k_0 - k_n)$.

It can also be shown that the integral of $D^t(z, \omega)$ around the cuts in the complex plane is identically zero, the contribution from any one term always being cancelled by that from its mate.

The solution in the region $d<z<\infty$ proceeds along essentially the same lines. The disturbance is identically zero until $(z-d)/u_{02}<t$ and reduces to the steady state solution when $z/u_{01}<t$.

When $u_{01} < 0$, so that one of the beams is flowing in the opposite direction, the method of solution given above breaks down, because it is no longer possible to assume that the modulation on the beams at the plane $z = 0$ is independent of the initial spatial distribution of disturbance. For the case of the infinite stream the easiest way out of the difficulty is probably to apply a Lorentz transformation and use a frame of reference with respect to which the beams are both flowing in the same direction. The solution given above can then be taken over unchanged.

For the bounded stream the solution will depend on the actual method by which the beams are injected into the interaction space. If the reverse beam is injected at the boundary $z = d$ from a separate cathode it would be logical to assume that $\rho_1(d, t)$ and $v_1(d, t)$ were independent of the initial spatial distribution of disturbance and also of $\rho_2(0, t)$, $v_2(0, t)$. We should then express $\rho_1^t(0, \omega)$ and $v_1^t(0, \omega)$ in terms of the independent modulations before carrying out the inverse transformation from the ω-plane into the t-plane. The resulting expressions for $\rho_1^t(0, \omega)$, $v_1^t(0, \omega)$ are complicated functions of the various partial waves $\exp(ik_\lambda z)$ with poles at the zeros of an expression of the general form

$$\sum_{\lambda=1}^{N} \sum_{\mu=1}^{N} [A_\lambda(\omega)B_\mu(\omega) - A_\mu(\omega)B_\lambda(\omega)] \exp i(k_\lambda + k_\mu)d = 0,$$

where k_λ is any one of the N roots of equations (A 5), (A 6), (A 7).

Dielectric Constant and Electron Density in a Gas Discharge

By R. E. B. MAKINSON,* P. C. THONEMANN,† R. B. KING and J. V. RAMSAY‡
School of Physics, University of Sydney

Communicated by H. S. W. Massey; MS. received 1st *March* 1951

ABSTRACT. Measurements have been made of the impedance at frequencies of the order 1,000 Mc/s. of a coaxial transmission line terminating in a small electrode immersed in the plasma of a low pressure discharge in mercury vapour. The frequencies f chosen were in the neighbourhood of that for which the real part ϵ_1 of the dielectric constant of the space near the electrode, as given by $\epsilon_1 = 1 - ne^2/(\pi m f^2)$ where n is the local electron density, was near zero. The effect of making a small change in the size of the electrode was also studied. The electron density in the neighbourhood of the electrode was measured by Langmuir probes. A minimum in reactance and a maximum in resistance (when plotted against discharge current) were found to indicate $\epsilon_1 = 0$ near the electrode, in agreement with a qualitative argument given. The possible use of such impedance measurements as a means of measuring electron density is discussed.

§ 1. INTRODUCTION

The recent development and availability of equipment for electrical measurements at frequencies upwards from 500 Mc/s. has led to the suggestion that such measurements may be of value in the study of gas discharges. In particular, since the real part ϵ_1 of the dielectric constant of a medium containing n free electrons per cm^3 is a function of n and of the frequency f, namely $1-ne^2/(\pi mf^2)$ with the usual assumptions*, it should be possible to measure n and conversely to test the above relation by comparison with Langmuir probe measurements.

---

* Now at Physics Department, University College, London.
† Now at Clarendon Laboratory, Oxford.
‡ Now at Royal College of Science, London.

While it is easy to produce an effect on microwave propagation or on the capacity of a condenser by ionization of a gaseous medium, it is difficult to devise a suitable experimental arrangement to realize the above aims for the following reason. In a discharge tube the electron density varies with position in the tube; consequently if the interpretation of the high-frequency measurements in terms of density distribution is to be feasible, a configuration of particular simplicity is essential. The measured quantities should be affected mainly by the value of ϵ_1 near a small region over which n varies but little; further the electrode systems used must not affect n appreciably in their neighbourhood. Arrangements described previously (e.g. Imam and Khastgir 1937) have not met these requirements.

At a certain frequency, f_0 say, the value of ϵ_1 for a particular n will be zero. The methods here discussed aim at locating f_0 from special features to be expected in the electrical behaviour of a small electrode situated in or near a region where $\epsilon_1=0$.

In the method first proposed, the small electrode immersed in the discharge tube was made the end of a short transmission line which could be brought into resonance with an oscillator by means of an adjustable plunger. At fixed frequency (near 1,000 Mc/s.) a discharge current was sought at which a change in the size of the electrode was without effect on the setting of the plunger for resonance. Such a discharge current would then be that for which ϵ_1 vanished in the neighbourhood of the electrode at this particular frequency. The method failed on trial, since no such condition could in general be obtained. The connection of the disc D to the electrode E always appeared to add a positive capacity to the end of the resonant line, and a lack of sharp resonance made settings difficult in the region where a negative sign might be expected. This failure was ascribed to the effect of the additional loss introduced by the increase in size of the electrode (equivalent to a comparatively low resistance across the end of the line) altering the resonant frequency of the line.

A more successful procedure was to measure at various frequencies the complex impedance $Z=R+iX$ presented between the lead to the electrode and a screen surrounding the tube containing the discharge, as the discharge current was varied.

The behaviour of R and X when ϵ_1 passes through zero at the electrode is not readily calculable, but analogous calculations for electrically simpler electrode systems indicate clearly the behaviour to be expected, namely a sharp rise in R and a minimum in X as the region where $\epsilon_1=0$ passes into the space between the electrodes where the induction is concentrated. An estimation of electron density may thus be made which was found to agree reasonably well with Langmuir probe determinations.

§ 2. THE DISCHARGE TUBE

A discharge in mercury vapour was chosen for study as being convenient and already well studied. In Figure 1 is sketched the hot-cathode discharge tube with which most of the measurements were made; a pool-cathode tube was found to be insufficiently steady for the purpose.

---

* The validity of this relation for free electrons has been extensively discussed, for example, by Darwin (1934). Salpeter and Makinson (1949) have shown that the relation breaks down when the electrons in their random free motion traverse large distances (compared with those over which the field varies appreciably) during a cycle.

The tube was thoroughly outgassed before distilling over a small quantity of mercury M, firing the getters G and sealing off. The total emission of the cathode exceeded 6 amperes; most of the readings were made with anode currents of the order of 1 ampere . the potential drop in the tube being approximately 40 volts. The anode supply was from a 100-volt accumulator via a rheostat.

The electrode E was of tungsten 1 mm. in diameter, glass being fused over the whole of its length inside the tube (approximately 1·5 cm.) except for approximately 1 mm. at the tip. A nickel disc D, 5 mm. in diameter, could be placed in contact with E or removed to the other side of the tube by moving the glass rod and nickel armature R with the aid of a magnet. The Langmuir probes P were of tungsten wire 0·091 mm. in diameter, 12·7 mm. in length, situated at the same distance from the wall as the tip of the electrode E. The inner conductor of the coaxial cable connector F was solidly connected to E, the outer conductor being soldered to the copper screen S which surrounded the central portion of the tube between the probes. The temperature of the tube was not controlled.

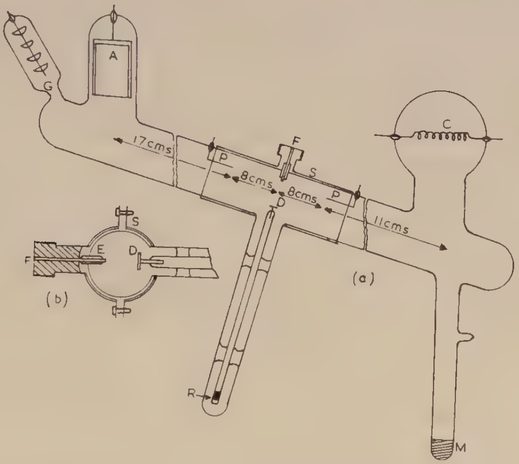


Figure 1. (*a*) The discharge tube. (*b*) Detail of electrode connection and screening. C oxide-coated cathode, A graphite anode, G getters, M mercury, E high frequency electrode, P probes, D nickel disc, R armature. Tube, Pyrex 2 in. nominal diameter.

§ 3. IMPEDANCE MEASUREMENTS

A low power oscillator employing a 'lighthouse' tube fitted with a calibrated attenuator and covering a range of frequencies 600 to 1,000 Mc/s. was used, together with a receiver with a similar frequency range in which the output was registered by a diode and milliammeter. A particular frequency f being selected, the fitting F was connected by coaxial cable to a sliding-probe slotted-line impedance measuring equipment. The latter was designed for use with 50-ohm cable, a length of which equal to an integral number of electrical half wavelengths was always used. The impedance $Z=R+iX$ presented by the terminal F was thus measured for a range of anode currents in the tube, i.e. for a range of electron densities at E. From the standing-wave ratio on the slotted line (measured with the attenuator) Z was computed with the aid of a Smith chart (Smith 1944). The electrode E was floating as far as its steady potential was concerned.

Curves showing the resistance R and reactance X so obtained for several frequencies are shown in Figures 2 and 3.

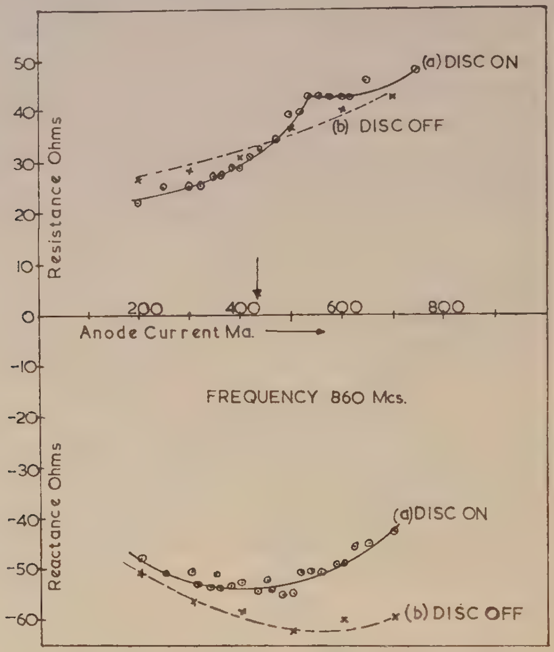


Figure 2. Resistive and reactive components of impedance of electrode system at 860 Mc/s (*a*) with disc D in contact with electrode E, (*b*) disc removed.

For a number of values of the anode current maintained in the course of these measurements the electron density was determined by the procedure of Langmuir (Langmuir 1923, Langmuir and Mott-Smith 1924, Mott-Smith and Langmuir 1926) from current potential curves of the probes P. The results from the two probes agreed closely and their mean was taken as giving the electron density near the tip of electrode E. For each frequency used, the value of the anode current at which the expression $1-\boldsymbol{n}\boldsymbol{e}^2/(\pi \boldsymbol{m}\boldsymbol{f}^2)$ vanished near the tip of E was thus located; these values are indicated by arrows on Figures 2 and 3 and their positions may be compared with the features of the curves for R and X.

§ 4. DISCUSSION

The electron density within the discharge tube near the electrodes may be supposed (Killian 1930) to be axially symmetrical with a maximum on the axis. The computation from such a distribution and the actual electrode configuration of curves directly comparable with those of Figures 2 and 3 is a mathematical problem clearly of some difficulty.

A much simpler basis for qualitative discussion may be sought in expressions for the impedance of a condenser with plane-parallel electrodes, in which the real part of the dielectric constant varies linearly from unity at one plate to an assumed value $\epsilon_1(t)=1-at$ less than unity (it may be negative) at the other plate. The distance t between the plates is supposed much less than a wavelength. We suppose that $\epsilon(x) = \epsilon_1(x) + i\epsilon_2(x)$ where $i\epsilon_2$ is an imaginary part representing dielectric losses (due to collisions) and we will suppose $\epsilon_2(x) = bx$.

We may now find the impedance of such a condenser as a circuit element by an elementary calculation, using the relations

$$E[\epsilon_1(x) + i\epsilon_2(x)] = D = 4\pi Q = 4\pi J/i\omega; \quad V = \int_0^t E\,dx,$$

where D is the electric induction and E the electric field at a point x in the dielectric, V the potential difference between the plates, Q the charge per unit area on the plates, J the current per unit area and ω the angular frequency. The values of $R/|X_0|$ and $X/|X_0|$, where $R + iX = Z = V/J$ and X_0 is the value of the reactance for vacuum as dielectric, are:

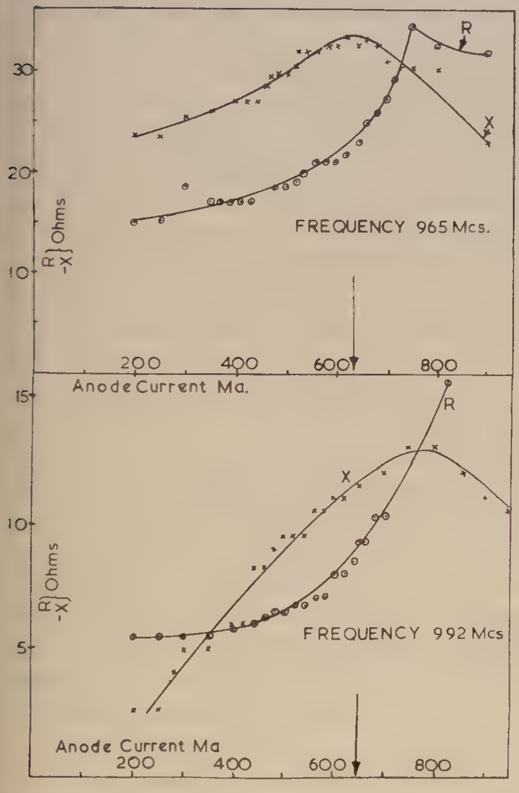


Figure 3. Resistive and reactive components of impedance of electrode system, at three frequencies, with disc D in contact with electrode E. The arrows indicate discharge currents for which ϵ_1 (computed from probe measurements) vanishes at D.

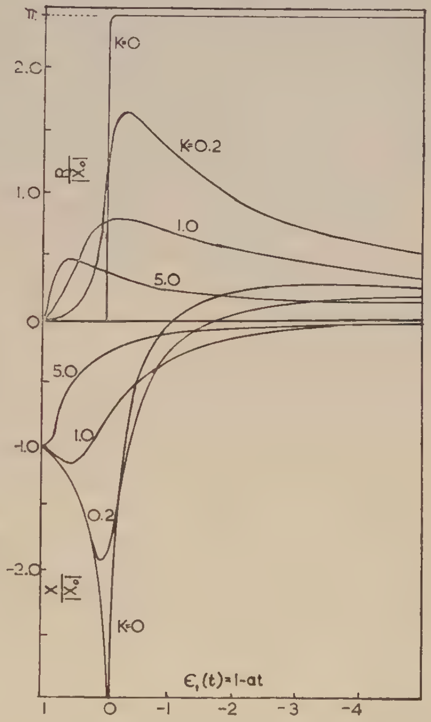


Figure 4. Resistive and reactive components calculated for a parallel plate condenser, with dielectric constant of medium varying linearly between the plates.

$$\frac{R}{|X_0|} = \frac{K}{2at(1+K^2)} \ln\{(1-at)^2 + K^2a^2t^2\} + \frac{1}{at(1+K^2)} \tan^{-1}\frac{Kat}{1-at},$$

$$\frac{X}{|X_0|} = \frac{1}{2at(1+K^2)} \ln\{(1-at)^2 + K^2a^2t^2\} - \frac{K}{at(1+K^2)} \tan^{-1}\frac{Kat}{1-at}.$$

These are shown in Figure 4 as functions of $\epsilon_1(t)$ for various values of $K = b/a$.

For very small values of b, corresponding to low collision losses in an electron-gas dielectric, it is seen that the resistive part of the impedance nevertheless rises abruptly to a finite value ($\pi|X_0|$, independent of b) when $\epsilon_1(t)$ is reduced through zero, so that there is an infinitesimally thick layer of dielectric in which $\epsilon_1 = 0$. This finite loss in an infinitesimal region arises, of course, because of the very high value attained by the electric vector $\mathbf{E}$ in that region. It will be noted also that $X/|X_0|$ has a minimum, when b is not large, occurring very nearly when ϵ_1 just vanishes at one electrode.

With electrodes of other shapes and less simple laws of variation of ϵ_1 with position, we may still expect an abrupt rise in R and a minimum in X whenever a transition of ϵ_1 through zero appears in a region between the electrodes where the induction is concentrated.

Now the curves of Figures 2 and 3 give clear evidence of such behaviour. The maximum or kink in R and minimum in X may be interpreted as arising when, with a general increase in electron density in the tube, the layer in which $\epsilon_1 = 0$ passes the disc D on the electrode E in moving nearer the wall. As it does so it moves into a region where most of the H.F. electric induction is concentrated. As in Figure 4, the minimum of X occurs first.

According to the Langmuir probe measurements of density, ϵ_1 would vanish at the disc, if there were no local reduction of density near the disc and supporting electrode, at the discharge currents marked, which lie close to the rather broad minima in X and well below the kinks in R. It is possible that the latter are displaced to the right by local reduction of density near the disc (which might be reduced by using a fine grid or spider in place of the disc D).

Further measurements, of the kind shown in Figure 3 with a range of pressures, gases and frequencies would, of course, be necessary to establish the procedure of locating either the minimum in X or the kink in R as a reliable method for exploring electron densities. It would, however, seem to offer some promise for this purpose as an alternative to available methods. For example it might prove possible to measure the rate of decay of electron density after short pulsed discharges by using a 'gated' receiver circuit on the impedance measuring equipment (or by modulating the oscillator) so that measurements refer only to a narrow interval of time at a known adjustable time after each pulse.

ACKNOWLEDGMENTS

The authors wish to thank Professor V. A. Bailey for his suggestions which led to this project, Dr. K. Landecker for a number of discussions, the Radiophysics Laboratory of the Council for Scientific and Industrial Research for the loan of essential equipment, and Mr. H. G. Phillips for skilled glass work. One of us (P. C. T.) held a Commonwealth Research Studentship while engaged on this work.

REFERENCES

Darwin, C. G., 1934, *Proc. Roy. Soc.* A, **146,** 17.
Imam, A., and Khastgir, S. R., 1937, *Phil. Mag.*, **23,** 858.
Killian, T. J., 1930, *Phys. Rev,*, **35,** 1238.
Langmuir, I., 1923, *Gen. Elect. Rev.*, **26,** 731.
Langmuir, I., and Mott-Smith, H., 1924, *Gen. Elect. Rev.*, **27,** 449, 538, 616, 762, 810.
Mott-Smith, H., and Langmuir, I., 1926, *Phys. Rev.*, **28,** 727.
Salpeter, E. E., and Makinson, R. E. B., 1949, *Proc. Phys. Soc.* B, **62,** 180.
Smith, P. H., 1944, *Electronics*, **17,** 130.

Conductivity, Hall Effect and Thermo-electric Power of Selenium Single Crystals

By K. W. PLESSNER*
Associated Electrical Industries Ltd., Aldermaston, Berks.

Communicated by E. Billig; *MS. received* 9*th January* 1951

ABSTRACT. Single crystals of selenium have been grown from the vapour phase and their conductivity and thermo-electric power were measured over a range of temperatures. The Hall effect was determined for one crystal. The conductivity obeys a relatively simple exponential law $\sigma=\sigma_0 \exp(-eE/kT)$, with $eE=0{\cdot}13$ ev. and the room temperature conductivity between 10^{-5} and $0{\cdot}5\times10^{-6}$ ohm^{-1} cm^{-1}. Both thermo-electric power and Hall effect rise slightly with increasing temperature, indicating a fall in the current carrier concentration. Values calculated from thermo-power and Hall effect agree reasonably well, and give about 10^{14} positive holes per cm^3. Because of an experimentally established correlation between thermo-power and conductivity, differences in conductivity between various crystals are ascribed to different hole concentrations. Deviations from Ohm's law, the rapid increase in mobility with temperature and its low absolute value (about $0{\cdot}14$ cm^2/v.sec. at room temperature) all point to the existence of potential barriers within the crystals. It is shown that a range of barriers of different heights could explain the observed behaviour. The conductivity is anisotropic (ratio about 10 : 1), whereas the thermo-power is not.

§ 1. INTRODUCTION

SELENIUM has been recognized as a semiconductor for more than 50 years and many of its properties were measured long ago. Yet to-day, when the conduction mechanism in most semiconductors is believed to be understood, an explanation of the behaviour of selenium is still outstanding. The present measurements were made in an attempt to clarify the situation a little by determining the concentration of free charge carriers in addition to the conductivity over a range of temperatures. Such a determination has in the past made possible a very complete interpretation of the behaviour of many semiconductors such as germanium and silicon.

The concentration of current carriers n (electrons or positive holes) is generally found by measuring the Hall coefficient $R(=3\pi/8ne)$; in theory, the thermo-electric power θ also gives a value for n. As far as is known, θ has only been used for an evaluation of n quite recently (Henkels 1950), and only once has a comparison between R and θ been made on the same sample, Ge, in order to test the formula for θ (Lark-Horovitz 1945). The agreement was good. The temperature dependence of θ has frequently been compared with that of the conductivity σ. If the conductivity obeys the usual semiconductor law $\sigma=\sigma_0 \exp(-eE/kT)$ then it is expected that θ will fall with temperature as E/T.

For a number of oxides this has been verified, but cuprous oxide, silicon carbide and some sulphides have shown a different behaviour, in that the thermo-electric power remains constant or even rises with increasing temperature. This has led to suspicion of the theory of the thermo-electric power (Greenwood and Anderson 1949).

* Now with British Dielectric Research Ltd., London W.12.

In selenium the Hall effect is very small* and has only been measured on one sample of halogenated microcrystalline material (Eckart and Kittel 1941). The thermo-electric power, however, is a fairly easily measured quantity and at the time of starting this work, two sets of results on microcrystalline selenium were available (Hochberg and Sominsky 1938, Borelius and Gullberg 1945); both sets showed θ increasing slightly with the temperature. The conductivity of microcrystalline selenium is, however, a complicated function of temperature and it was therefore decided to work with single crystals.

Previous measurements of the conductivity of single crystals of selenium (Müller 1938) showed a relatively simple variation with temperature and it was hoped that the thermo-electric power would give the corresponding n values. The measurements described below show that an interpretation is not as simple as was hoped and, in order to confirm the evaluation of n from θ, efforts were concentrated on measuring the Hall effect on one crystal. The main results of the experiments have already been communicated in short notes (Plessner 1950).

§ 2. EXPERIMENTAL METHOD

2.1. *Preparation of Crystals*

The method of growth from the vapour phase described by Brown (1914) has been used. This consists in allowing Se to sublime under a small pressure of inert gas to form crystals on the colder part of the sublimation vessel. The Se was placed at the bottom of a vertical glass tube about $1\frac{1}{4}$ in. diameter which stood in a heated brass cup so that the Se was kept molten at about 250° c. Crystals grew inwards and downwards from the wall of the tube in the region where it emerged from the heater. Most crystals had the external shape of prisms standing on a hexagonal base and were elongated in the direction of the prism axis, i.e. along the crystal c-axis. Nearly all these needles were hollow, only the biggest ones having appreciable solid portions. These were used for the measurements, the hollow portions being ground away until the prisms had flat end faces.

Rotating-crystal x-ray diffraction photographs showed single rows of well-defined spots, indicating single crystals. Small disorientations of parts of the crystal would, however, not have been detected.

Some crystals grew in the form of flat plates, the c-axis lying in the plane of the plate. One of these was used to measure the conductivity and thermo-electric power at right angles to the c-axis.

2.2. *Measurement of Conductivity and Thermo-electric Power*

The specimens, about 7 mm. long and 1·5 mm. in diameter, were clamped between spring-loaded brass blocks, which contained chromel–alumel thermocouples close to the contacting surfaces. One of the blocks could be heated by 5 to 10° c. with respect to the other so that a temperature gradient could be set up in the specimen. The thermal E.M.F. was determined by measuring the voltage between the two chromel wires, so that the thermo-electric power found was that of the system chromel–selenium–chromel. The conductivity was measured by passing a known current through the specimen and measuring the potential between two probes of pointed gold wire making contact at distances about one third and two thirds along the specimen. The specimen current was found by measuring the voltage across a known series resistance.

---

* Although the Hall constant R has a normal value, the current which can be passed through the specimen is limited by the low conductivity and also by noise. In terms of the applied voltage, the Hall voltage is directly proportional to the mobility, and this is very small in Se.

All voltages were measured with the aid of a valve electrometer used as null detector in a potentiometer-type circuit. The known voltage, which was placed in opposition to the unknown, was derived from a standard Muirhead potentiometer for values up to 2 v. and from an H.T. battery with voltmeter for higher values. Instability of the electrometer valve set the lower limit of measurable voltages at 0·1 mv. For temperature measurement the thermo-E.M.F. generated by the chromel–alumel couples was measured using a galvanometer as null detector, the sensitivity in this case corresponding to 0·1° c. The accuracy of the thermo-electric power measurement was therefore within approximately 3 %, using a temperature difference of 5° c.

The accuracy of the conductivity measurement was of a similar order, being determined by voltage readings on a meter and measurements of the dimensions of the specimen and of the probe separation.

During the measurements the specimen was kept in darkness to avoid photo-effects. The screening can could be electrically heated, or cooled by immersion in a CO_2–acetone bath, so that the specimen temperature could be varied between $-65°$ and $+210°$ c.

2.3. *Hall Effect Measurement*

Attempts were made to determine the Hall voltage in a conventional manner by means of D.C. current and field. However, the voltage fluctuations appearing at the Hall probes when a current was passed through the specimen were so large as to mask any Hall voltage. It is known that semiconductor noise has a spectrum which falls off with increasing frequency (noise voltage $\propto 1/\sqrt{f}$) and therefore an A.C. method of measurement was used. An alternating Hall voltage may be obtained either by alternating specimen current or by alternating magnetic field. In practice t was found that the noise voltage was larger when passing an alternating current than when passing a direct current through the specimen, and therefore the latter method was used. The specimen assembly was placed in a field alternating at 30 c/s. and the Hall voltage was fed into conventional amplifiers via cathode followers and differential amplifiers (long-tailed pair), which enabled the difference n voltage between the two probes to be measured. A homodyne detector allowed a very narrow bandwidth to be employed, this being determined by the time-constant in the measuring instrument circuit (about 2 sec.). The homodyne being phase-sensitive, the signal voltage in phase with H could be distinguished from nduced voltages which lagged 90° behind H. Also, the Hall voltages caused by opposite polarities of specimen current, which differ 180° in phase, gave deflections n opposite directions on the homodyne meter. In this way, the deflection was wice that which would have been obtained with a simple rectifier. The sensitivity was variable from 10μv. to 10 mv. full-scale deflection.

Unwanted signals were caused by electromagnetic induction in any loop ormed by the probe circuit, the induced voltage lagging 90° in phase behind the ield. This pick-up was reduced below significance by varying the loop area until t was algebraically zero. However, when this was done (by shifting one of the eads from the probes relative to the other), a voltage in phase with H remained, the cause of which could not be traced. This voltage threw some doubt on the results obtained, since the signal voltage might merely have been a modulation of this residual pick-up. However, the reproducibility and the magnitude of the effect over a temperature range indicated that the measured voltage was the true Hall effect. The accuracy of measurement was limited by noise fluctuations and by fluctuations in pick-up voltages to about ±25%.

Some further details of the apparatus might be of interest. Contact to the crystal was made by evaporating nickel electrodes on the end faces (current electrodes) and in the form of spots on the sides (Hall electrodes)*. These electrodes were touched by molybdenum springs and tungsten whiskers respectively. The electrode assembly was screened by copper sheet against electrostatic pick-up, but the copper was cut away over the area of the pole-pieces to avoid eddy currents. This portion was, however, further screened by a film of nichrome evaporated on mica. The surface resistivity of the films was about 40 ohms, and by passing a current along them the specimen could be heated to any desired temperature. The temperature was measured by means of a chromel-alumel thermocouple attached to one of the current electrodes.

§ 3. RESULTS

The conductivity measurements confirm Müller's work, which has only been extended to a small degree by measuring the conductivity up to higher fields. As has been stressed in a previous note (Plessner 1950) selenium crystals show departures from Ohm's law at quite low fields even when potential probes are used. For sufficiently high fields (>200 v/cm.) the conductivity always rises with increasing field, the more markedly at low temperatures, but with smaller fields above room temperature the conductivity first falls. A typical set of curves is given in Figure 1.

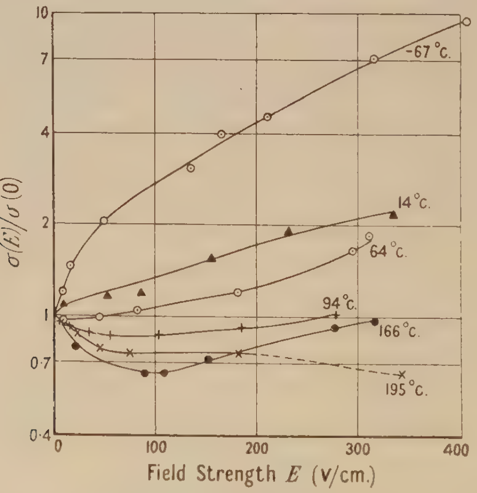


Figure 1. Conductivity against applied field, normalized to agree at zero field (parallel to c-axis).

Some further results in agreement with Müller's conductivity measurement are briefly described in §§ 3.1, 3.2 and 3.3.

---

* Nickel forms practically no barrier when in contact with selenium.

3.1. *Creep Effects*

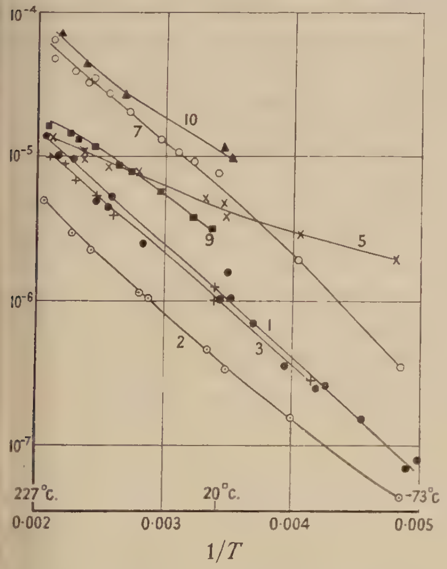
The application of a high field changed the resistance as measured with a lower field. The changes were of different magnitude and sign at different temperatures and no correlation with other effects could be found. The specimen generally recovered within a few minutes. Müller obtained some curves of resistance plotted against time which showed variations of several per cent, but these have not been repeated, since the potentiometric method used was not suitable.

3.2. *Annealing*

On first heating a fresh crystal an irreversible increase in resistance takes place, but this resistance then remains stable provided the highest temperature is not again surpassed. When heating to just below the melting point the resistance (measured at room temperature) rises by a factor of about 10; after several months the resistance falls to its original value. The temperature coefficient of the resistivity near room temperature is not changed. The cause of this effect has remained obscure, although a number of tests were made to find it. Thus fresh crystals were heated up with and without the application of an electric field, in and out of contact with metal electrodes, in air and in a vacuum of about 10^{-5} mm. Hg, and with and without illumination. In each case the increase in resistance took place. Rapid cooling had no effect on the annealing. X-ray analysis proved inconclusive, since single crystal rotation photographs both before and after annealing showed slightly elongated spots due to the size of beam and crystal. A more sensitive technique may throw some light on the matter.

3.3. *Conductivity as a Function of Temperature*

Figure 2 shows curves of log σ plotted against inverse temperature for a number of crystals, over the temperature range $-65°$ to $+210°$ c. The lines are fairly straight and, with one exception, nearly parallel. Müller gives similar results, his



igure 2. Conductivity parallel to the c-axis for several Se crystals plotted against 1/temperature.

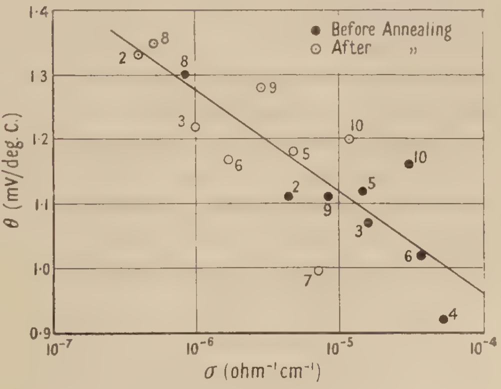


Figure 3. Thermo-electric power against conductivity for several Se crystals at room temperature.

lines extending to -180° c. While the absolute values of conductivity at room temperature vary between 10^{-5} and $0{\cdot}5\times10^{-6}$, the activation energy [$\boldsymbol{e}E$ in the expression $\sigma=\sigma_0\exp(-\boldsymbol{e}E/\boldsymbol{k}T)$] is nearly constant at $0{\cdot}13\pm0{\cdot}02$ ev. It is demonstrated in the discussion (§4) that the exponential law is not due to a process such as the liberation of current carriers from impurity centres, and therefore the more usual form $\sigma=\sigma_0\exp(-\boldsymbol{e}E/2\boldsymbol{k}T)$ is not used.

The measurements shown in Figure 2 were made at low field strengths with current flow in the direction of the crystal *c*-axis. All the crystals were prepared under similar conditions except No. 9, which was grown using crystals from previous sublimations as raw material, and No. 10 which was grown under a higher vacuum in a region of constant temperature. There is no indication why crystals 1–7 behaved so differently.

3.4. *Thermo-electric Power*

The sign of the thermo-electric effect indicates conduction by positive holes. Room-temperature values of the thermo-electric power θ are contained in Figure 3, where they have been plotted against the corresponding conductivity both before and after annealing. The small numbers refer to the individual identity of the crystals. It is seen that annealing always raises θ as well as the resistivity.

The variation of θ with temperature was also measured for each crystal, but the amount of scatter makes it unprofitable to present curves of θ against T. It may be said that for most crystals θ rose slightly with temperature in the range 20° to 210° c., at most by $0{\cdot}24$ mv/° c. This is in agreement with Henkels' (1950) measurements on a single crystal. Below room temperature measurements on the crystals with high resistance were difficult, but the results obtained showed that remained nearly constant also at these temperatures.

3.5. *Hall Effect*

The Hall effect has been measured on crystal No. 5, which was particularly large so that the evaporated electrodes could be easily applied; it also had a low resistivity and correspondingly low noise level.

A typical set of results is given to indicate the orders of magnitude involved:

Temperature (° c.)	20	191	Magnetic field (r.m.s.) (gauss)	4,800
Specimen current (μa.)	45	12	Length of crystal (cm.)	0·65
Specimen voltage (v.)	34	20	Width of crystal (cm.)	0·2
Hall voltage (μv. r.m.s.)	100	350	Cross section (cm^2)	0·0305

From these data the conductivity σ, the current carrier concentration n and the carrier mobility v can be calculated, using well-known formulae.

In Figure 4, $\log\sigma$, $\log v$ and $\log n$ have been plotted against $1/T$. It is seen that n is nearly constant and that the variation of $\log v$ as well as of $\log\sigma$ is linear.

3.6. *Anisotropy*

The conductivity $\sigma_\perp$ at right angles to the *c*-axis has been measured on only two crystals, but shows good agreement for these. The various effects mentioned above also apply to $\sigma_\perp$: deviations from Ohm's law are similar below room temperature, but less pronounced above room temperature, creep effects have been noticed and annealing occurs, but to a lesser extent. The dependence of $\sigma_\perp$ on temperature is shown in Figure 5, where for crystal No. 7 the values of $\sigma_\parallel$ parallel to the *c*-axis have also been included. It is seen that the ratio $\sigma_\parallel/\sigma_\perp$ is about 10, the slopes of the curves being similar above room temperature.

The thermo-electric power with the temperature gradient perpendicular to the *c*-axis was 1·35 mv/° c. for crystal No. 8 and 1·02 for No. 7. Crystal No. 7 gave 0·99 mv/° c. along the *c*-axis, the difference of 3% being hardly significant. These values did not vary appreciably with temperature.

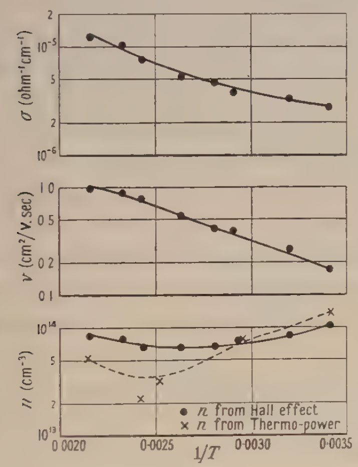


Figure 4. Conductivity σ, mobility v, and carrier concentration n, against 1/temperature for crystal No. 5 along *c*-axis.

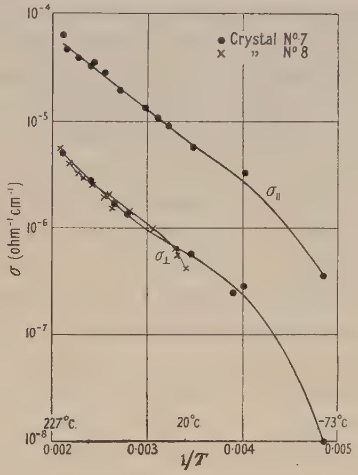


Figure 5. Conductivity along and perpendicular to *c*-axis against 1/temperature.

§ 4. DISCUSSION

In order to classify Se among the known semiconductors it is necessary to decide whether conduction is intrinsic or extrinsic. Intrinsic conduction, when electrons are thrown up from the full into the empty energy band, is generally characterized by only very little scatter of conductivity values and by a relatively steep slope of the $(\log \sigma, 1/T)$ curve (equivalent to half the width of the forbidden band). The results obtained have neither of these characteristics. Moreover, according to recent work by Becker and Fan (1950) on Si and Ge, the width of the forbidden band may be estimated from the long-wave limit of the fundamental lattice absorption. Since the absorption edge of Se crystals lies at about 7,000 A.* the slope of the $(\log \sigma, 1/T)$ lines should be $\frac{1}{2} \times 1{\cdot}8$ ev. for intrinsic conduction, whereas only 0·13 ev. is observed. Hence conduction must be extrinsic, i.e. due to current carriers supplied by impurity centres. In this case also, an exponential relation between σ and $1/T$ is generally obtained due to the temperature-activated liberation of current carriers, the mobility varying only slowly with temperature. On the basis of the conductivity results alone, this might be expected to apply here. However, the theoretical relation between the thermo-electric power θ and the current carrier concentration n is $\theta = (k/\boldsymbol{e})(2 + \log N_A - \log n)$ wher $N_A = 2\{2\pi \boldsymbol{mkT}/\boldsymbol{h}^2\}^{3/2}$ so that the constant or slightly increasing θ which has been observed, implies a constant or decreasing n. Thus, the impurities appear to be all ionized at the temperature of the experiments. This may also be inferred from another consideration. Using the room temperature value of θ only, a carrier concentration of about $10^{14}\,\mathrm{cm}^{-3}$ is obtained and with an activation energy of $2\times0{\cdot}13$ ev. (height of acceptor centres above full band) ionization would be virtually complete at all temperatures above -140° c.

---

* Information kindly supplied to the author by his colleague Mr. J. J. Dowd.

With constant n and the relation $\log\sigma\propto-1/T$ it is seen that for Se crystals $\log v\propto-1/T$, a relation difficult to reconcile with any theory of the mobility v. Impurity scattering, giving $v\propto T^{3/2}$, would lead to an increase by a factor of only 2·3 between room temperature and the melting point, whereas actually a factor of about 10 is found. This contradiction, due to the use of the above formula for θ, has been found by other workers on various materials (Greenwood and Anderson 1949) and has led to suspicion of the validity of the formula. However, as Figure 4 shows, the values of n derived from measurement of the Hall effect on Se crystal No. 5 agree well with those derived from the thermo-electric power (the crosses denote values of n derived from θ) and therefore confirm the behaviour outlined.

Another somewhat indirect confirmation of the proportionality between θ and $(C-\log n)$ is presented in Figure 3, which shows a definite correlation between the room-temperature conductivity and the thermo-electric power of the various crystals. There is a linear relationship between $\log\sigma$ and θ, the correlation coefficient being 0·84, and the slope of the best straight line not far from $\boldsymbol{k}/\boldsymbol{e}\,(=86\mu\mathrm{v}/^\circ\mathrm{c.})$. Since $\sigma=n\boldsymbol{e}v$, theory predicts such a correlation if v is the same for all the crystals. The actual value obtained is $v=0{\cdot}14\,\mathrm{cm}^2/\mathrm{v.\,sec.}$

The experimental results may be summarized by the following three facts: Ohm's law is disobeyed at fields exceeding about 5 v/cm., the concentration of charge carriers is approximately independent of temperature, and the conductivity and the mobility increase with temperature as $\exp(-\boldsymbol{e}E/\boldsymbol{k}T)$ where $\boldsymbol{e}E\approx0{\cdot}13$ ev. Apparently the only way to explain these facts is to postulate internal barriers. The resistance to current flow is then concentrated in the barriers, giving a low apparent mobility whose temperature dependence is connected with the properties of the barriers rather than with those of the crystal lattice. Also, intergranular barriers generally give rise to deviations from Ohm's law.

There are several possible types of high-resistance contacts between grains: (*a*) The area of contact between grains is small. Such a 'spreading resistance' does not give a temperature variation of boundary resistance differing from that of the grain resistance. (*b*) The grains are separated by a potential barrier due to an intervening layer of foreign, or at least structurally different, material. A thin potential barrier can be crossed by tunnel effect, whilst a thick barrier, if not too high, can be overcome by thermal excitation (thermal barrier). Only the latter type gives rise to a resistance varying as $\exp(\boldsymbol{e}E/\boldsymbol{k}T)$, as is required to explain the temperature dependence of the mobility.

In the presence of barriers it is not certain that either the thermo-electric power or the Hall effect gives a correct measure of the carrier concentration in the grains. As pointed out by Henisch (1948), however, the thermo-electric power is measured without any current flowing through the specimen and should therefore not be influenced by high-resistance barriers. Similarly, the transverse field in the Hall effect is measured potentiometrically and should not be affected by barriers. Since the bulk of the material consists of ordered domains and the current flowing through them is correctly measured, the value of n derived from Hall voltage and current density should be the true carrier concentration within the grains. A more detailed analysis by Volger (1950) leads to the same conclusion.

Sufficient is known about potential barriers to allow a rough calculation of their resistance to be made. The resistance per unit area of a barrier of thickness d and constant height E is $d\exp(\boldsymbol{e}E/\boldsymbol{k}T)/n\boldsymbol{e}v$. For barriers of other shapes the exponential term contains the maximum barrier height, but for the same thickness the resistance is lower. If there are z barriers per centimetre then the resistivity is $\rho=(zd\,n\boldsymbol{e}v)\exp(\boldsymbol{e}E/\boldsymbol{k}T)$. The apparent mobility $1/\rho n\boldsymbol{e}=(v/zd)\exp(-\boldsymbol{e}E/\boldsymbol{k}T)$ is thus exponentially dependent on the temperature, in agreement with the experimental evidence. Substitution of typical experimental data gives $v/zd=20$. Even when a value as low as 2 cm^2/v.sec. is assumed for v, corresponding to a mean free path of about the lattice spacing, one obtains $zd=0{\cdot}1$, i.e. the fraction of the crystal consisting of barriers is one-tenth. Such a high fraction should, of course, be visible under the microscope, but no irregularities have been found. This argument shows that thermal barriers of the correct height to reproduce the temperature variation give too great an absolute value of the conductivity. By allowing the barrier height to increase with temperature ($E=E_0+T\partial E/\partial T$) a better approximation may be obtained. For instance, if $\partial E/\partial T=2\times10^{-4}\,V/^\circ$ c. then $zd=0{\cdot}01$ or 1%. Such an assumption, however, appears rather artificial and it is simpler to postulate a range of barriers of different heights acting in parallel. Most of the current would then pass through the lowest barriers, giving a temperature dependence with low activation energy, while the absolute value of the conductivity obtained with this model is small because of the reduced area of cross section through which the current flows. Further, if some of the high barriers are so thin that the tunnel effect is operative, then the strong variation with temperature of the deviations from Ohm's law may be explained. Transmission by the tunnel effect is independent of temperature but increases rapidly with field. Therefore at low temperatures, when the conductance of the low thermal barriers is small, the tunnel effect makes a relatively greater contribution to the total current and the drop of resistance with increasing field is more pronounced than at high temperatures, where the current through thermal barriers predominates. It must be assumed that the thermal barriers have a resistance which rises with voltage, like a Schottky or Mott rectifying barrier in the blocking direction. (For a diagram showing the deviations from Ohm's law at various temperatures see Figure 1.)

The non-uniform current distribution just postulated will affect the Hall coefficient to some extent. Non-uniformity in the direction of the Hall field is of no significance since both Hall voltages and current densities add up linearly, but non-uniformity in the direction of the magnetic field will have some effect, since in this case regions of different Hall voltage are in parallel. It may be expected that the average taken by a Hall probe of fairly large area will correspond roughly with the average current density.

The number of current carrying barriers per centimetre length which are effectively in series may be estimated from the field at which non-linearity begins—about 5 v/cm. If this corresponds to $\boldsymbol{k}T/\boldsymbol{e}$ or 25 mv. per barrier, there would be 200 barriers per centimetre. The physical nature of the barriers is obscure. As has been mentioned, some of the crystals are of very regular appearance and there seems no reason why any discontinuities large enough to give potential barriers should occur during crystal growth.

Since in annealing the drop in conductivity is accompanied by a rise in thermo-electric power, the effect appears to consist of a decrease in the number of positive holes and acceptor centres, while the mobility remains constant. This is consistent with the fact that the temperature coefficient of the conductivity remains unchanged.

As the thermo-electric power is nearly the same in the two crystal directions, the anisotropy in conductivity must be due to the mobility being ten times higher along than across the *c*-axis. The small difference in thermo-electric power which has been found may be connected with a difference in effective mass, which appears as $(\boldsymbol{k}/\boldsymbol{e}) \log \boldsymbol{m}^{3/2}$ in the formula.

The nature of the acceptor centres unfortunately remains doubtful. It is reasonably certain that there are about 10^{14} per cm^3 and that they are all ionized even below room temperature. From this it may be deduced that the energy levels of the centres are not more than 0·24 ev. above the full band, but no further deductions can be made from the results obtained.

Se crystals have been grown (Dauber 1946) in the presence of halogen vapours, without the conductivity showing any increase. Conversely, our attempt at growing crystals from purer raw material did not result in a lower conductivity. It seems likely, therefore, that the acceptor centres do not consist of ‘ impurities ’ in the form of foreign atoms, but only of imperfections in the Se crystal lattice. This view is also supported by the fact that in micro-crystalline Se the carrier concentration depends markedly on heat treatment, but scarcely at all on halogen content.

§ 5. CONCLUSION

The conclusion is reached that the Se crystals grown from the vapour phase are not true single crystals but contain internal barriers. Müller arrived at the same conclusion on the evidence of non-ohmic behaviour, and Henkels deduced that a crystal grown from the melt was more truly single because of its nearly ohmic property. The present evidence for barriers consists, however, not only of deviations from Ohm’s law, but also of the values of Hall effect and thermo-electric power and their variation with temperature. Henkels’ thermo-electric measurements agree well with the present results, and it is therefore probable that his crystal grown from the melt also contained internal barriers.

A slightly wider conclusion may perhaps be drawn from the correlation of Hall effect and thermo-electric power. Se is an example of those semiconductors in which the thermo-electric power remains nearly independent of temperature while the conductivity shows the usual exponential increase. The variation of the Hall effect with temperature shows that the constant thermo-electric power should be interpreted as indicating constant carrier concentration. It seems natural to extend this interpretation to other semiconductors having constant thermo-electric power and to explain the change in their conductivity in terms of intergranular barriers, and not in terms of the excitation of current carriers into the conduction band, as is usually done.

ACKNOWLEDGMENTS

The author is indebted to Dr. E. Billig for discussions and advice, to Mr. W. P. Melling for help with the electronic equipment, and to Dr. T. E. Allibone for permission to publish this paper.

REFERENCES

Becker, M., and Fan, H. Y., 1949, *Phys. Rev.*, **76,** 1530, 1531.
Borelius, G., and Gullberg, K., 1945, *Ark. Mat. Astr. Fys.*, **31a,** 10.
Brown, F. C., 1914, *Phys. Rev.*, **4,** 85.
Dauber, H., 1946, *FIAT Final Report* No. 706.
Eckart, F., and Kittel, A., 1941, *Naturwissenschaften*, **29,** 371.
Greenwood, N. N., and Anderson, J. S., 1949, *Nature, Lond.*, **164,** 346.
Henkels, H. W., 1950, *Phys. Rev.*, **77,** 734.
Henisch, H. K., 1948, *Elect. Comm.*, **25,** 163.
Hochberg, B. M., and Sominsky, M. J., 1938, *Phys. Z. Sowjet*, **13,** 198.
Lark-Horovitz, K., 1945, *NDRC*, 14–585.
Müller, T., 1938, *Sitz.Ber. phys. med. Soz. Erlangen*, **70,** 7.
Plessner, K. W., 1950, *Nature, Lond.*, **165,** 970, **166,** 1073.
Volger, J., 1950, *Phys. Rev.*, **79,** 727.

Conductivity and Hall Effect of Micro-crystalline Selenium containing Iodine Impurities

By K. W. PLESSNER*
Associated Electrical Industries Ltd., Aldermaston, Berks.

Communicated by E. Billig; MS. received 9th January 1951

ABSTRACT. The conductivity and Hall effect of micro-crystalline selenium have been measured as a function of iodine content under two different conditions of crystallization, over a range of temperatures. The conductivity increases with iodine concentration, a maximum being reached at a concentration of about 0·03 atomic per cent. The Hall effect shows this increase to be due to an increasing mobility, the carrier concentration remaining nearly constant. The effect of carrying out the first crystallization of the amorphous Se at 175° c. instead of 110° c. is to reduce the conductivity by a factor of ten. This is due to reductions in both mobility and current carrier concentration. With increasing temperature the mobility rises and the carrier concentration falls, the product of the two resulting in a rising conductivity with samples crystallized at 175° and a slightly falling conductivity with samples crystallized at 110° c. The results are discussed in terms of the grain structure of the material. Microscopic examination has shown that the grain size is different for different crystallization temperatures and it is suggested that larger crystals contain a smaller density of acceptor centres. The fall in carrier concentration with increasing temperature is thought to be due to the relief of stress set up due to anisotropic contraction when the material is cooled from the annealing temperature. The low absolute value of the mobility and its rise with temperature are ascribed to intergranular high resistance barriers.

§1. INTRODUCTION

A large amount of experimental work has been carried out on selenium ever since its practical applications were discovered, but no great understanding of the conduction mechanism has been reached. It is realized that a source of complications is the micro-crystallinity of selenium, and from time to time attempts have been made to use single crystals for a clearer understanding of the conduction mechanism in the selenium crystal lattice. Such attempts have not clarified the situation greatly, the conclusion generally being reached that even ‘ single ’ crystals contain internal barriers which dominate the resistivity. The ease of preparing samples of micro-crystalline selenium in comparison with single crystals and its greater importance from the point of view of applications have led to a number of recent researches on this material.

---

* Now with British Dielectric Research Ltd., London W.12.

The most important electrical measurement to be made on a semiconductor, apart from the conductivity, is that of the Hall effect, since it allows a calculation of the density (and the sign) of current carriers. In selenium the Hall effect is very small and has escaped detection so far, except for a measurement of one sample of doped selenium by Eckart and Kittel (1941). It has only recently been realized that the thermo-electric power also gives a value for the carrier density n, at least theoretically. Henkels (1950 a) was the first to calculate n from the thermo-electric power measured on micro-crystalline Se, and his results are particularly relevant to the measurements described below. Measurements of the conductivity of micro-crystalline Se containing known amounts of halogens have recently been made by Blackburn (1948) and Schweikert (1950). In the present work conductivity measurements on halogenated Se have been supplemented by a determination of the carrier concentration, using the more conventional Hall method. This provides a useful check on the calculation of n from the thermo-power. The particular object of the present measurements has been to determine to what extent the carrier concentration in micro-crystalline Se is affected by (*a*) the addition of a halogen impurity, (*b*) the conditions of crystallization of the Se and (*c*) the temperature.

§ 2. PREPARATION OF SPECIMENS

Selenium containing from 0·1 to 0·0003 atomic per cent of iodine was prepared as follows. Using the purest Se available (Johnson Matthey spectroscopically standardized Se shot) some SeI_4 was made by melting together Se and Analar grade iodine. A 0·1% mix was then prepared by diluting the SeI_4. To avoid the small amount of iodide evaporating* before the Se was molten and the mixing could be completed, the pellet of SeI_4 was wrapped into a piece of amorphous Se made plastic by slight warming. The mix was heated in a Pyrex test-tube and was only kept above the melting point for about ½ to 1 minute to allow vigorous stirring. It was then chilled. Further concentrations were made by diluting the 0·1% mix, this time merely melting, stirring for a short time and chilling. The amorphous Se resulting from the quenching was then crystallized and shaped into discs ¾ in. in diameter and about 0·5 mm. thick under pressure in a heated steel mould. A piece of amorphous Se was dropped into the mould maintained at a constant temperature, and as the Se went through the plastic phase (40° – 80° c.) it was pressed into shape by the upper piston. Two crystallizing temperatures were used, 110° and 175° c.; samples prepared at these temperatures are referred to as belonging to groups A and B respectively. In order to obtain smooth discs at 110° c. a moulding pressure of 50 kg/cm^2 was applied for one minute followed by 19 minutes at 10 kg/cm^2, while at 175° c. the pressure was maintained at 10 kg/cm^2 for 20 minutes, because the higher pressure produced cracked specimens. It was estimated that the Se would assume the temperature of the mould in a few seconds. At 110° c. crystallization appears to progress sufficiently slowly for the Se to have reached this nominal temperature before crystallization has gone far, but at 175° C. the transformation is so rapid that most of the Se is probably crystalline before this temperature is reached. In addition to this effect, the heat evolved in the phase change will tend to warm up the material (Borelius 1949).

After crystallization all the specimens were given a uniform annealing treatment for 40 minutes at 218° c. in air. In analogy with Schweikert's measurements

---

* SeI_4 disintegrates at 100° c., iodine melts at 114° c. and boils at 183° c.

(1950) this should result in maximum conductivity without much loss of halogen. The chemical analysis of his samples showed a loss of about 70% Cl or 30% Br content after 40 minutes at 210° c. The vapour pressure of I_2 being much lower, the loss by evaporation is not expected to exceed about 10%, even though our shape of specimen has a slightly larger ratio of surface area to volume. Moreover, Schweikert found that the fractional loss was the same for all his concentrations.

Nickel electrodes were evaporated on to the discs in the shape shown in Figure 1. No appreciable barrier layer is expected between Ni and Se, and tests on a rectifier made with an evaporated Ni counter-electrode showed that the contact resistance is less than 20 ohm cm². With the area of each main electrode shown in Figure 1 about 0·2 cm², the series resistance due to the contacts may be as much as 2 × 100 ohms. Probe measurements on a sample with a particularly low resistance* (5,000 ohms) showed that the voltage drop between the electrode and the Se just next to it was only 4% of the applied voltage. Calculation shows that about 2% of this is due to the fact that the electrodes are applied to the faces and not the rims of the discs. Thus the contact resistance was of no great importance.

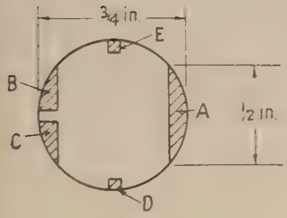


Figure 1.

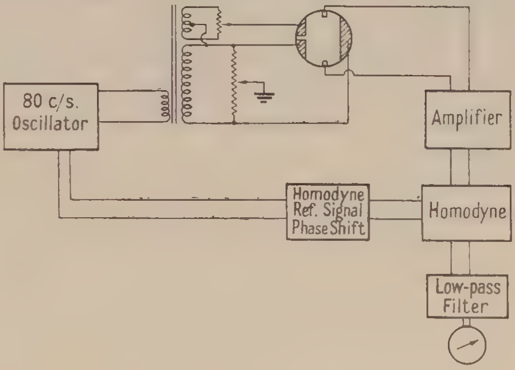


Figure 2.

§ 3. ELECTRICAL MEASUREMENTS

3.1. *Conductivity*

The resistance of the specimen was measured by finding the current on application of a small D.C. potential (2 V.) between the electrodes A and B–C in parallel. Care was taken to exclude light. As the resistance was found to increase slowly for a period of days after the final heat treatment, all the measurements were made several days after this treatment.

Owing to the circular shape of the samples, chosen for convenience in moulding, the conductivity could not be calculated directly from the dimensions and the resistance. A reasonable estimate can, however, be made by integration. The result obtained in this way was confirmed to within 1% by testing one specimen first in its circular form and after cutting it to a rectangular shape.

3.2. *Hall Effect*

As there is excessive D.C. current noise an A.C. method was used to measure the Hall effect. A block diagram of the circuit is shown in Figure 2. An 80 c/s. current was passed through the specimen by applying about 10 V. r.m.s. to the electrodes A and B–C, and the Hall voltage appearing at the electrodes D and E when a D.C. magnetic field was switched on was amplified and measured. Because of unavoidable inhomogeneity in the specimen and misalignment of the electrodes an alternating voltage appeared at the Hall terminals even without the magnetic field being applied. This was balanced out by feeding slightly different voltages on to the two halves B and C of the split current electrode. Thus the current flow was made asymmetrical to compensate for resistance dissymmetry. Further, a type of Wagner earth arrangement was used to ensure that each Hall electrode was nearly at earth potential. This consisted of a potentiometer with an earthed tapping across the current terminals on the specimen. To adjust for balance the voltage of only one Hall electrode with respect to earth was amplified and this was then reduced to zero by varying the potentiometer tapping. On some particularly inhomogeneous specimens, when it was possible to balance the Wagner earth but not the specimen as a whole, the Hall effect was measured using only the one electrode. It was verified with a normal specimen that one half the usual Hall voltage was measured in this way.

---

* The bulk resistance should increase more than the contact resistance, but this has not been tested.

Amplifier design was conventional, detection being performed by a homodyne circuit. This ensures a low bandwidth ($=1/2\pi\tau$, where τ is the time constant of the measuring instrument and circuit) and also gives a measurement which is phase-sensitive. This is convenient in so far as it allows the Hall effect to be distinguished from a magneto-resistive effect, since the former changes sign with the magnetic field, whereas the latter does not.

The shape of the specimen does not permit a direct calculation of the Hall constant from the Hall voltage for two reasons. Firstly, the lines of current flow are not straight, secondly, the current electrodes have an appreciable short-circuiting effect on the Hall probes, since they are relatively close to them. Both these points have been treated by Isenberg, Russel and Greene (1948). Regarding the non-uniform current distribution it was shown that this is of no importance, the Hall voltage depending only on the total current. The shorting effect of the end electrodes was, however, shown to introduce an error if the ratio of length to width is less than three. In our case, where this ratio is about unity, the measured Hall voltage V_H' is only 70% of the maximum possible value obtained from the formula $V_H = RIH/t$, where R is the Hall constant, I the total current, and t the specimen thickness in the direction of the magnetic field H.

Consequently, the value of the Hall constant was obtained as $R = V_H' \, t/0{\cdot}7\, IH$. The current carrier concentration n was calculated from $R = 3\pi/8ne$ and the mobility v from $v = \sigma/ne$.

The accuracy of measurement varied from sample to sample. It was determined mainly by noise fluctuations, but sometimes by a high rate of drift of the balance point. For a given applied voltage, the Hall voltage is determined by the mobility v, and so, as a rule, the greater v, the greater the accuracy. As a general indication, ±20% may be quoted as the limits.

§ 4. EXPERIMENTAL RESULTS

4.1. *Room Temperature Measurements*

All the results obtained at room temperature are contained in Figure 3 plotted with iodine concentration as abscissa. The uppermost curve, giving the conductivity of samples crystallized at 110° C., agrees well with Schweikert's results on chlorinated Se in the range between 10^{-4} and 10^{-5} parts of halogen but deviates at higher iodine contents. A similar drop in conductivity at higher concentrations has been observed by Blackburn (1948). The conductivity of samples crystallized at 175° c. follows a similar curve, but is almost ten times lower.

The current carrier concentrations n derived from the Hall effect are plotted below σ in Figure 3. The relatively greater scatter in these values is mainly due to experimental error. Values of n for both sets of samples show a tendency to increase with iodine concentration, but those for group B are appreciably lower than those for the crystallization temperature of 110° c. The mobility v determined from the values of σ and n is also included in Figure 3.

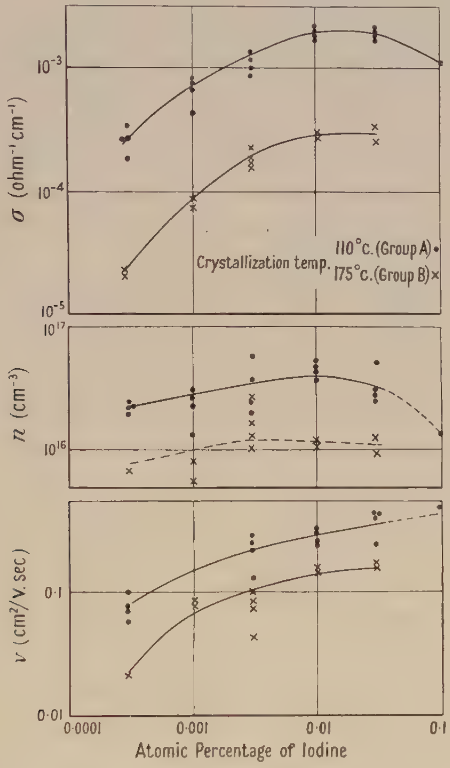


Figure 3. Conductivity σ, carrier concentration n, and mobility v, of micro-crystalline Se against iodine content.

4.2. *Measurements at Variable Temperature*

Conductivity and Hall effect have been measured at different temperatures for six specimens—three from each set of samples with constant crystallization temperature. The temperature ranged from room temperature to 200° c.

Curves of σ, n and v plotted logarithmically against $1/T$ are given for four of these specimens in Figure 4–7. It is seen that the two specimens which were crystallized at 110° c. have an almost constant conductivity up to about 100° c., after which the conductivity *falls* with increasing temperature. The specimens crystallized at 175° c., on the other hand, show a rising conductivity over the whole range, the slope of the line corresponding to 0·11 ev. in the case of the specimen containing 0·0003% I_2, and approximately 0·06 ev. for the specimen with 0·03% I_2 (E in $\sigma \propto \exp -E/\boldsymbol{k}T$). Considerable hysteresis is generally observed, but in view of Blackburn's statement that even after extremely long exposures to each temperature the hysteresis did not disappear, no more than $\frac{1}{2}$ to 1 hour was allowed for each temperature. Curves of ($\log n$, $1/T$) show that n decreases with

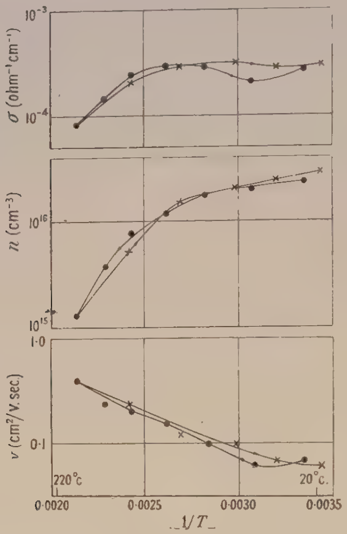


Figure 4. Conductivity σ, carrier concentration n, and mobility v, against temperature for a group A specimen; 0·0003% iodine.

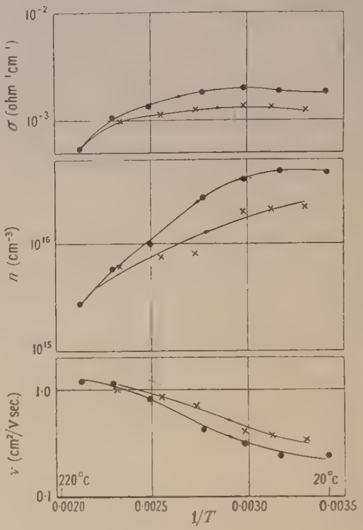


Figure 5. σ, n and v against temperature for a group A specimen, 0·03% iodine.

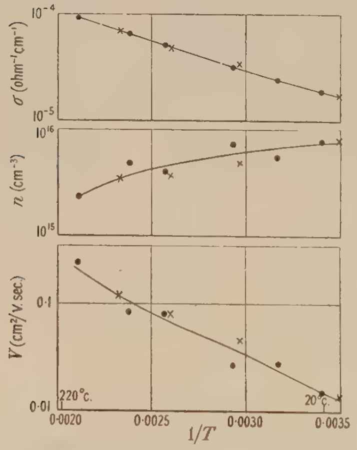


Figure 6. σ, n and v against temperature for a group B specimen, 0·0003% iodine.

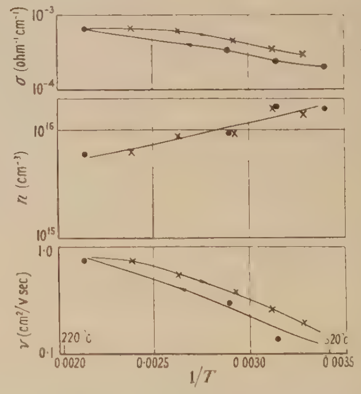


Figure 7. σ, n and v against temperature for a group B specimen, 0·03% iodine.

increasing temperature in all cases, but more markedly so for the specimens crystallized at 110° c. This may be considered as the cause of the falling conductivity. The mobility values derived from n and σ have again been included for convenience in discussion.

4.3. *Density*

The specimens crystallized at 175° c. (group B) showed slight porosity, and to check this the density of eight samples of group B was compared with that of eight samples of group A. The density of a single crystal of selenium is 4·78 (Straumanis 1940) and that of amorphous Se 4·15 (Borelius and Paulson 1946). The densities obtained for crystallization temperatures of 110° c. and 175° c. were 4·43 and 4·16 ± 0·03, i.e. 93% and 87% of the single crystal densities, respectively.

The 110° c. specimens showed no porosity even under microscopic examination. Nevertheless, the amount of amorphous Se which would have to be specified to explain the observed density as a mixture of amorphous and crystalline material is far too large to be plausible (about 55%), and no doubt some porosity does exist.

4.4. *Microscopic Examination*

In order to investigate the effect of different crystallization temperatures on the structure of the Se, some photomicrographs were taken with vertical illumination. The surfaces of the samples show outlines of what are presumably crystallites, and a significant difference exists between the samples crystallized at different temperatures. Figures 8 (*a*) and (*b*) (Plate*) show typical surfaces of samples crystallized at 110° and 175° c. respectively, each containing 0·001% I_2. Clearly the crystal size is much greater with the higher crystallization temperature. A sample of group B containing 0·03% I_2 had a still larger crystal size. No structure within the crystals is visible, probably because the surface has been moulded to conform with the polished steel piston. Greater detail can be seen by examining the surface obtained on breaking a specimen in two. Figures 8 (*c*) and (*d*) show such views of specimens containing 0·0003% I_2 and crystallized at 110° c. and 175° c. respectively. In spite of the poor depth of focus, the striations typical of cleaved crystals of Se can be clearly seen, especially in (*d*). All the specimens have undergone the 218° c. annealing treatment as well as the crystallizing treatment. The size of the crystals may be estimated very roughly from these photographs. In selenium treated at 110° c. they are 2–3μ, in selenium crystallized at 175° c. about 10μ long.

The fact that larger crystals are obtained at the higher temperature is easily understood, both from Borelius' work (1949) and from general considerations. The crystal size is determined by a balance between the rate of nucleation and the rate of crystal growth. In the range of temperatures of interest the rate of nucleation falls with increasing temperature while the rate of crystal growth increases (owing to the reduced viscosity). Both factors favour a larger crystal size at the higher temperature. Spherulites have frequently been found in group B material, a particularly clear one being reproduced in Figure 8 (*e*). In selenium crystallized at 110° c. spherulites have been observed only before the 218° c. treatment, not afterwards. This may be due to strong grain growth breaking up the many small spherulites formed at 110° c., while the large spherulites formed at 175° c. are not so easily broken up by further grain growth. Borelius' statement that no spherulitic crystallization takes place above 130° c. is not necessarily in contradiction to the present observations, since in the experimental technique used here crystallization is almost certainly initiated well below the nominal temperature of 175° c.

The principal conclusion to be drawn from the microscopic examinations is that the crystal size is about three times larger in material of group A than that of group B.

---

* For Plates see end of issue.

§ 5. DISCUSSION

There is very little doubt that the resistance in Se is determined largely by intercrystalline boundaries. At least three compelling arguments may be cited to support this contention: non-linearity of the current–voltage curve at high fields (Henniger 1938), low values of mobility (Henkels 1950 a, from thermo-electric power, and the present paper) and radio-frequency measurements (Dauber 1946).

In the presence of barriers the significance of Hall measurements may be questioned. The Hall field set up across crystallites is a function of the current density and although the longitudinal field in the crystallites is reduced by the presence of barriers, the current density is measured correctly. Therefore, the voltages across the crystals give true indications of the carrier density *n* inside them and it only remains to show that the total Hall voltage measured across the sample is the sum of all the individual voltages in spite of the intervening barriers. These barriers, which are in parallel with the crystallites as far as the main current flow is concerned, take very little current but have a high Hall constant (small *n*). Provided the mobility in the barriers is not far different from that in the crystal, the barriers contribute Hall voltages according to their thicknesses and so the Hall field is constant through both crystals and barriers. Therefore, the Hall measurement gives a correct value for the current carrier density *n* within the crystal grains (see also Volger 1950).

Inspection of Figure 3 shows that the values of *n* are affected far more by the crystallization temperature of the specimens than by their iodine content. This fact makes it probable that the current carriers in Se are not produced by halogen impurity but are due to lattice defects, which are connected with the micro-crystalline structure of the material. Microscopic evidence has shown the large difference in crystal size between group A and B, and it appears that the larger the crystals, the fewer are the current carriers. In the extreme case of ' single ' crystals the carrier density is only about $10^{14}\,\mathrm{cm}^{-3}$ (Henkels 1950 a, Plessner 1951). The small effect which the iodine content seems to have on *n* may be of a secondary nature, since it is known that halogen impurity increases the rate of crystallization of Se.

Turning now to the temperature dependence of *n*, Figures 4–7, it is seen that *n* always falls with increasing temperature. It does so more strongly in group A specimens, where it has the higher values at room temperature, than in group B specimens. As far as the author is aware only one other case of a Hall constant rising (*n* falling) with temperature has been reported, viz. germanium at very low temperatures (Hung 1950). The explanation given in that case unfortunately does not apply here, because it requires that the conductivity should be independent of temperature, which is not the case here. The only explanation which can be offered is of a very qualitative nature. Krebs (1949) has shown that x-ray diffraction lines of micro-crystalline Se sharpen as the temperature is increased and broaden as it is decreased. This effect is reversible, and is, of course, contrary to normal experience, where the broadening is due to thermal vibrations which increase with temperature. Further, at room temperature, the line broadening is greater in samples crystallized at low temperatures than in those prepared at high temperatures. These observations suggest that line broadening and the density of current carriers are related, since they both increase together. Krebs has shown that the x-ray line broadening is due to lattice distortion, and suggests that part of this distortion is due to faulty crystallization, e.g. Se chains with both right- and left-handed screw directions may be built into the same crystal, and part of it to thermal stresses set up by unequal contraction of the crystals and the intercrystalline amorphous material. In a randomly orientated mass of Se crystals, another source of large thermal stresses is the anisotropic contraction of the crystals ($\alpha = -18 \times 10^{-6}$ along the hexagonal *c*-axis and $+74 \times 10^{-6}$ along the *a*-axis (Straumanis 1940)). The greater mechanical strength is to be expected along the *c*-axis owing to valence binding and therefore on cooling from the annealing temperature, at which no stresses exist, the expansion of the *c*-axis will prevent the *a*-axes of neighbouring crystals from contracting freely and will cause a range of enlarged values of *a* to be registered. It is not clear how the distortion of crystals gives rise to current carriers, but it is interesting to note that at the annealing temperature all the specimens have nearly the same value of *n*.

In order to explain the observed conductivities, the effect of intercrystalline boundaries must be considered. The boundary material probably consists of amorphous selenium, whose resistivity at room temperature is generally quoted as approximately 10^{12} ohm cm. Extrapolation of Henkels' measurements on liquid selenium indicates an even higher value (Henkels 1950 b). At the junction between the crystalline and amorphous selenium a step in the potential will occur whose height depends on the relative work functions of the two materials and on possible space charges near the boundary. The amorphous selenium thus forms a potential barrier over which the current carriers have to be excited. A potential well is, of course, equally possible, but would not give a high resistance.

The effect of a potential barrier depends on its height and thickness, and current carriers may overcome a barrier by two mechanisms—thermal excitation or the tunnel effect. In either case the current passed per barrier will be proportional to the number of positive holes available, i.e. to *n*. Therefore, the conductance of the barriers must be proportional to the experimentally determined quantity σ/n or σ/ne, i.e. v. It is seen (Figures 4–7) that v always increases with temperature by a factor of 5 to 15 between room temperature and 200° C. It is fairly certain in view of this strong temperature dependence that the barriers are of the low and thick type, which are overcome by thermal excitation. An argument already developed in a companion paper (Plessner 1951) on selenium single crystals leads to the conclusion that barriers of the right height to give the observed temperature dependence will give much too low a resistance, if their thickness is assumed to be a reasonable fraction of the grain thickness (not more than 5%). However, if a range of barrier heights exists, then most of the current will flow through the lowest barriers, which will thus dominate the temperature dependence, while the presence of the high barriers increases the resistivity.

The effect of iodine in increasing the apparent mobility, and with it the conductivity, may consist in a reduction in either the thickness or the height or the number of barriers. The last is unlikely, since an increase in crystal size, and therefore a reduction in the number barriers, in going from group A to group B is accompanied by a decrease in mobility. Porosity may account for some of this decrease but would hardly reverse the position. A reduction in height of the barriers should make the mobility less dependent on temperature, but no such trend is noticeable; it may be that the accuracy of the results is not sufficiently high to allow a decision on this point. If iodine acts by reducing the barrier thickness, explanation may be found in the process of crystallization. It is known that halogen accelerates crystallization by shortening selenium chains and prising open rings. In the same way it may allow neighbouring crystals to approach more closely by breaking up the intertwined selenium chains which arrest growth and constitute the barrier material. An electrical effect of the iodine on the barrier shape in the form of a space charge is also possible, but knowledge of the effect of halogens on the properties of amorphous selenium is insufficient to allow any speculation.

Finally, Henkels' results (1950 a) should be discussed in relation to those given above. The only type of specimen which can be compared directly is that crystallized by Henkels at 210° c. since all the others were crystallized at lower temperatures and were not given the 218° c. annealing treatment which all our samples received. The figures for that particular sample agree best with the present group A specimen containing 0·0003% iodine (Figure 4):

	σ		n		v	
Temp.	Henkels	Gp. A	Henkels	Gp. A	Henkels	Gp. A
50° c.	$1\cdot4 \times 10^{-4}$	2×10^{-4}	$9\cdot9 \times 10^{15}$	2×10^{16}	0·09	0·07
100° c.	$1\cdot5 \times 10^{-4}$	3×10^{-4}	$4\cdot8 \times 10^{15}$	$1\cdot5 \times 10^{16}$	0·2	0·15
190° c.	$0\cdot64 \times 10^{-4}$	$0\cdot8 \times 10^{-4}$	10^{15}	$1\cdot3 \times 10^{15}$	0·4	0·4

§ 6. CONCLUSION

It is found that the effect of iodine in raising the conductivity of micro-crystalline Se consists in a reduction in the intergranular resistances, while the number of current carriers in the grains is hardly affected. On the other hand, a change in the conditions of crystallization of the Se affects the number of current carriers and it is therefore concluded that the origin of the current carriers is closely related to the perfection of the crystallites and is not due to iodine impurity. This is further supported by the unusual dependence of the carrier concentration on temperature, viz. a decrease with increasing temperature, which agrees with x-ray diffraction evidence (Krebs 1949) of greater crystal perfection at higher temperatures.

The variation with temperature of the conductivity of micro-crystalline Se is determined by a sensitive balance between the mobility, which increases, and the carrier concentration, which decreases, with increasing temperature.

ACKNOWLEDGMENTS

The author wishes to express his thanks to Dr. E. Billig for frequent advice, to Mr. W. P. Melling for designing the electronic equipment, and to Dr. T. E. Allibone for permission to publish this paper.

REFERENCES

Blackburn, W. E., 1948, *J. Appl. Phys.*, **19,** 50.

Borelius, G., 1949, *Ark. Fys.*, **1,** 305.

Borelius, G., and Paulson, K. A., 1946, *Ark. Mat. Astr. Fys.*, **33a,** No. 7.

Dauber, H., 1946, *FIAT Final Report* No. 706.

Eckart, F., and Kittel, A., 1941, *Naturwissenschaften*, **29,** 371.

Henkels, H. W., 1950 a, *Phys. Rev.*, **77,** 734; 1950 b, *J. Appl. Phys.*, **21,** 725.

Henninger, F. P., 1938, *Ann. Phys., Lpz.* (5), **31,** 342.

Hung, C. S., 1950, *Phys. Rev.*, **79,** 727.

Isenberg, I., Russel, B. R., and Greene, R. F., 1948, *Rev. Sci. Instrum.*, **19,** 685.

Krebs, H., 1949, *Z. Phys.*, **126,** 769.

Plessner, K. W., 1951, *Proc. Phys. Soc.* B, **64,** 671.

Schweikert, H., 1950, *Z. Phys.*, **128,** 47.

Straumanis, M., 1940, *Z. Kristallogr.*, **102,** 482.

Volger, J., 1950, *Phys. Rev.*, **79,** 1023.

Some Conduction Properties of the Oxides of Cadmium and Nickel*

By C. A. HOGARTH
Physics Department, University of Reading

MS. received 23*rd February* 1951

ABSTRACT. The electrical conductivity and thermo-electric power of specimens of cadmium oxide and nickel oxide are measured as functions of temperature and of oxygen pressure, and the results are found to agree in general form with theories previously proposed by C. Wagner and by the present author. The results are used to verify Wagner's concept of a differential thermo-electric circuit, and a band model for CdO is proposed.

§ 1. INTRODUCTION

Many of the experimental results and values quoted in the literature for the electrical properties of semiconducting metallic oxides are concerned with pressed powders or sintered specimens (e.g. Gudden 1934). Only in recent years have experiments on single crystal oxide specimens been carried out (e.g. Sproull and Tyler 1951). These results are of much greater value than the previous since the effects of inter-granular barriers in powder specimens may be such as to vitiate completely the value of experimental data on such properties as electrical conductivity and Hall coefficient. The thermo-electric power is much less likely to be affected by the polycrystalline nature of the specimens, and the effects of inhomogeneities and inter-granular boundaries on this property have been considered in some detail by Henisch (1949). The work of Wagner and his collaborators, carried out on polycrystalline oxide specimens with a varying pressure of the surrounding oxygen, nevertheless provided valuable data on the conductive processes in metallic oxides. Although the absolute values quoted for the conductivity are not those which would characterize single crystals of the same materials, the variation with oxygen pressure of the electrical conductivity at constant temperature provides information about the conduction and about the relative concentrations of free charge carriers. By means of an argument based on chemical thermodynamics and the law of mass action, the concentration n_f of free charge carriers, electrons or holes, is shown to be given by

$$n_f = KP_{O_2}^{\mp 1/n} \qquad \ldots\ldots(1)$$

where n is a constant independent of temperature in Wagner's treatment, the negative sign refers to excess and the positive to deficit semiconductors, K is the equilibrium constant, and P_{O_2} is the oxygen pressure. Wagner made the assumption that since at any given temperature, electron concentration and conductivity are directly proportional, equation (1) would lead to a variation of conductivity of the form

$$\log \sigma = \mp \frac{1}{n} \log P_{O_2} + \text{const.} \qquad \ldots\ldots(2)$$

Wagner's work and that of other experimenters has confirmed the general form of equation (2) for several oxides, although the values of the constant n obtained do not always agree with the mass action value. The present author has shown (Hogarth 1948 a) how the use of equation (1) together with the usual treatment of the theory of thermo-electric effects provides equations of the form

---

* Part of a thesis submitted to the University of London for the degree of Ph.D.

$$\frac{dE}{dT} = -\frac{\boldsymbol{k}}{|\boldsymbol{e}|n}\log P_{O_2} \mp \text{const.} \qquad \ldots\ldots(3)$$

where dE/dT is the thermo-electric power defined as negative for excess and positive for deficit semiconductors, $\boldsymbol{k}$ and $\boldsymbol{e}$ have their usual significance, and the negative sign for the constant applies to excess and the positive sign to deficit semiconductors. The application of equation (3) to the spot readings of Wagner and his co-workers leads to values of the constant n in better agreement with the mass action values (Hogarth 1948 b) than those obtained from conductivity measurements.

The form of equation (3) has been verified for the deficit-conducting CuI by Nagel and Wagner (1934), but the only oxide so far investigated in this connection is CdO, for which Hogarth and Andrews (1949) found agreement in the general form of the equation, but lack of agreement in the numerical constants. No systematic investigation has hitherto been recorded for the variation of thermo-electric power with oxygen pressure for a deficit-conducting metallic oxide.

The variation of Hall constant with oxygen pressure had not hitherto been measured experimentally for any oxides, although the allied problem of the variation of Hall coefficient R of CuI with iodine vapour pressure was investigated by Bädeker (1909, 1911) who found that log R was proportional to log P_{I_2}. The present author has shown (Hogarth 1948 a) that the theoretical variation in Hall effect with oxygen pressure is given by

$$\log|R| = \pm\frac{1}{n}\log P_{O_2} \pm \text{const.} \qquad \ldots\ldots(4)$$

where the positive sign applies to excess and the negative sign to deficit semiconductors. A note on the variation of Hall constant with oxygen pressure for CdO has been published elsewhere (Hogarth 1951).

The previous experimental work of Hogarth and Andrews (1949) on the thermo-electric power of CdO was performed on cylindrical specimens 8 cm. long with one end maintained at room temperature and the other gradually heated through several hundred degrees. In the present work the results are confirmed by a more conventional method of measurement, and the variation of electrical conductivity and thermo-electric power with oxygen pressure are determined both for CdO and for the deficit-conducting NiO. Reference is made to some results on the Hall coefficient of CdO and its variation with oxygen pressure, and the nature of the conduction mechanism for CdO is discussed with special reference to the work of Wright (1949) and Meyer (1944).

§ 2. EXPERIMENTAL ARRANGEMENTS

(i) *Nickel Oxide—Thermo-electric Power*

Experiments similar to those of Hogarth and Andrews were carried out for various specimens of deficit-conducting NiO.

(ii) *Cadmium and Nickel Oxides—Electrical Conductivity and Thermo-electric Power*

The specimens used were cylindrical samples 0·93 cm. in diameter and approximately 1 cm. long made respectively from the purest obtainable CdO and NiO. These were formed by compression in a steel die at a pressure of 7 tons in^2 and were each of about 70% of the crystal density as determined by x-ray methods. The specimen could be held securely between platinum disc electrodes and mica insulators in a steel clamp. By this means contact resistances were reduced to a minimum. Current and potential leads of platinum were fitted, and chromel–alumel thermocouples were soldered to each electrode disc. The whole assembly was enclosed in a vitreosil tube permanently sealed at one end and fitted with vacuum seals at the other, which could be lowered into an electric furnace. By adjustment of the position of the tube in the furnace a uniform temperature for the specimen could be arranged or a small temperature difference (about 25°) maintained between its ends. The general arrangement was similar to that of Anderson and Morton (1945). The electrical measurements were made with the use of a Tinsley thermo-electric potentiometer and an associated switching circuit. Electrical conductivity and thermo-electric power could easily be measured, the usual precautions for reversal of current when measuring conductivity being taken. The vacuum arrangements were substantially as described in a previous paper (Hogarth and Andrews 1949). The experiments with CdO were carried out from room temperature to about 550° and with black NiO over the same range, but with green NiO only from 200–550° c.

§ 3. RESULTS OF EXPERIMENTS

(i) *Black NiO—Thermo-electric Power*

Specimens of black NiO (shown by Prasad and Tendulkar (1931) to contain about 0·5% excess oxygen) were examined by the method of Hogarth and Andrews (1949). Thermo-E.M.F. was measured as a function of temperature difference with one end of the specimen maintained at approximately room temperature. From the slopes of the resulting curves the thermo-electric powers were calculated and found to decrease slowly with increasing temperature. The variation of thermo-electric power with oxygen pressure is shown in Figure 1. The mean value of the constant *n* in equation (3) deduced from the curves of Figure 1 is 20, which is considerably larger than the value of 5 deduced from

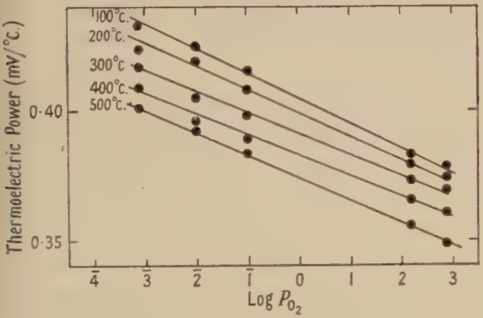


Figure 1. The relation between thermo-electric power and oxygen pressure at various temperatures for black nickel oxide.

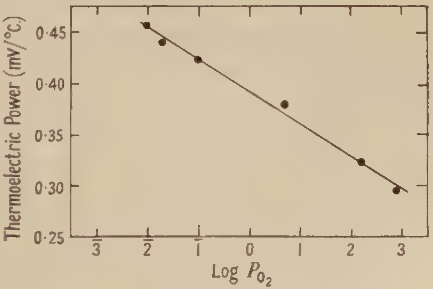


Figure 2. The relation between thermo-electric power and oxygen pressure at 400° c. for green nickel oxide.

Wagner's experimental results (Hogarth 1948 b), but is nevertheless nearer the theoretical value than the results already obtained by this method for CdO. The activation energy for black NiO deduced from conductivity measurements in the above temperature range was 0·55 ev. and the room temperature conductivity was 3×10^{-5} ohm^{-1} cm^{-1}.*

(ii) *Green NiO—Thermo-electric Power*

Specimens of green NiO were examined by the same technique. The variation of thermo-electric power with oxygen pressure at 400° c. is shown in Figure 2 and yields a value of 6·3 for the constant n, in excellent agreement with Wagner's predicted value. The activation energy was found to be 0·93 ev. and the conductivity at 400° c. was $1{\cdot}6 \times 10^{-2}$ ohm^{-1} cm^{-1}.

(iii) *CdO—Variation of Resistance with Temperature*

Figure 3 shows the variation of resistance of a CdO specimen with temperature. It was found that the resistance increases slowly with temperature up to about 150° c. and then decreases sharply. When the logarithm of the resistance is plotted against inverse temperature it is found that the high temperature end of the curve is linear, yielding an activation energy of 0·40 ev. Similar results were found with several specimens, yielding a mean value of 0·4 ev. for the activation energy.

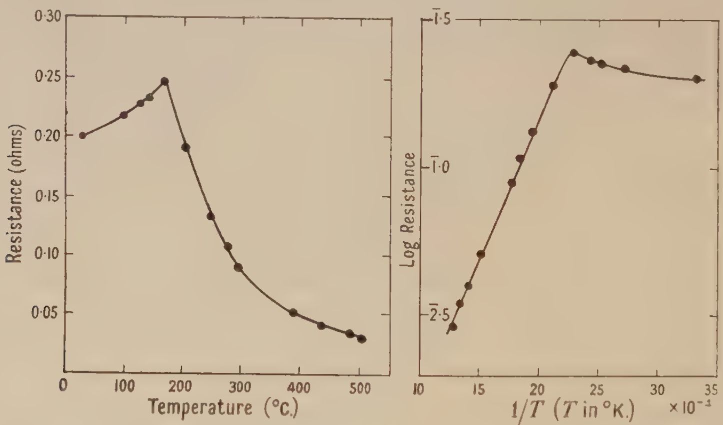


Figure 3. The resistance of a CdO specimen as a function of temperature.

(iv) *CdO—Oxygen Pressure Experiments*

The results of experiments on the electrical conductivity and thermo-electric power of a typical CdO specimen, using the method described in § 2 (ii) are given in Figures 4 (*a*) and 4 (*b*) for the temperature interval 200 to 550° c. It may be noted that the conductivity increases and thermo-electric power decreases as the oxygen pressure is decreased, in qualitative agreement with theory. When plotted in accordance with equations (2) and (3) the curves of Figures 4 (*c*) and 4 (*d*) are obtained and verify the form of the equations for an excess semiconductor.

The variation of the above properties with oxygen pressure and temperature within the limits described is accurately reproducible. After changing the air or oxygen pressure equilibrium values of the electrical constants were always reached within two minutes. In the curves of Figure 4 and subsequently, the effect of air as the surrounding gas is always plotted as its partial pressure of oxygen.

---

* The activation energy $\Delta\epsilon$ is defined by the equation $\sigma = A \exp(-\Delta\epsilon/2kT)$ relating conductivity and absolute temperature.

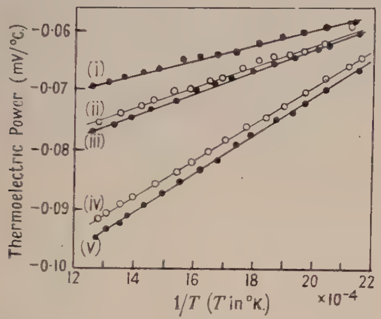


Figure 4 (*a*). The thermo-electric power of CdO as a function of temperature for various air and oxygen pressures.

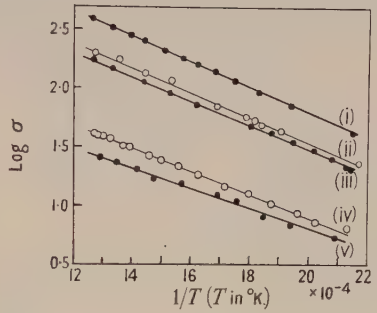


Figure 4 (*b*). The electrical conductivity of CdO as a function of temperature for various air and oxygen pressures.

(i) 0·001 mm. air, (ii) 0·01 mm. oxygen, (iii) 0·1 mm. air, (iv) 772 mm. air, (v) 777 mm. oxygen.

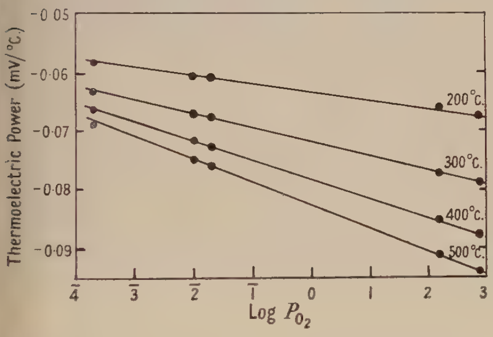


Figure 4 (*c*). The thermo-electric power of CdO as a function of oxygen pressure.

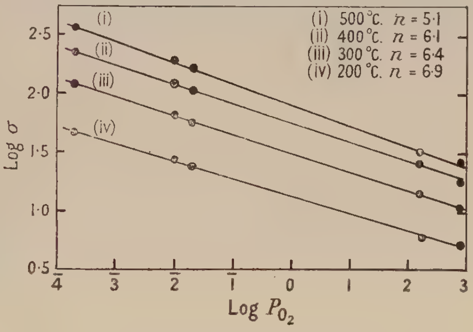


Figure 4 (*d*). The variation of conductivity of CdO with oxygen pressure.

(v) *Green NiO—Oxygen Pressure Experiments*

Results of measurements on the temperature variation of electrical conductivity and thermo-electric power for a typical small specimen are given in Figures 5 (*a*) and 5 (*b*). The results of these measurements are re-plotted as

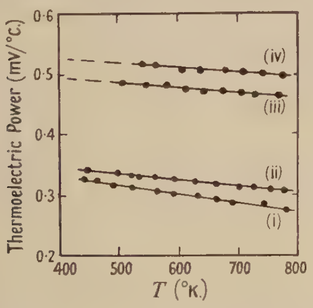


Figure 5 (*a*). The thermo-electric power as a function of temperature for green nickel oxide at different air and oxygen pressures.

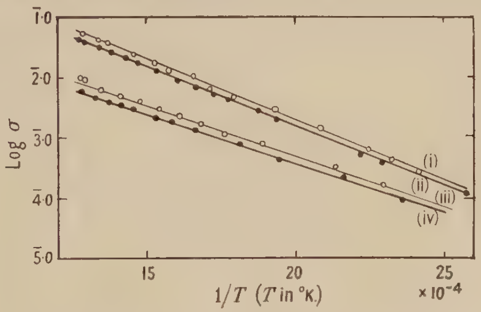


Figure 5 (*b*). The conductivity of green nickel oxide as a function of temperature for various air and oxygen pressures.

(i) 767 mm. oxygen, (ii) 768 mm. air, (iii) 0·01 mm. oxygen, (iv) 0·005 mm. air.

functions of oxygen pressure in Figures 5 (*c*) and (*d*) where again the general form of equations (2) and (3) is confirmed for a deficit semiconductor.

Again the results were reproducible although with NiO the time taken for equilibrium to be reached after a change of oxygen pressure was sometimes as much as two hours.

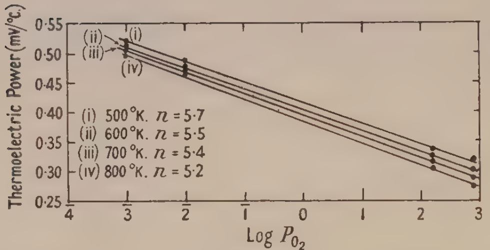


Figure 5 (*c*). The thermo-electric power of green nickel oxide as a function of oxygen pressure.

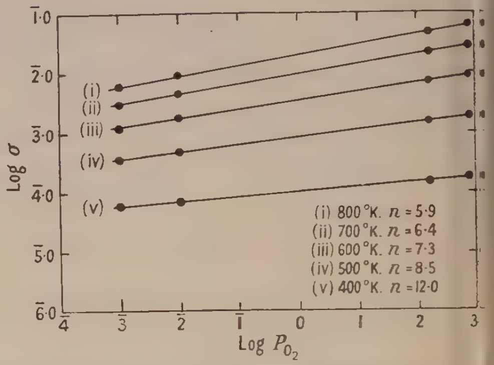


Figure 5 (*d*). The conductivity of green nickel oxide as a function of oxygen pressure.

§4. DISCUSSION OF RESULTS

(i) *Nickel Oxide*

The variation of electrical conductivity and thermo-electric power with oxygen pressure is in accordance with equations (2) and (3) and the experimental results may thus be considered to verify them. If values of the constant n are calculated from each set of results at various temperatures the values for the specimen whose properties are summarized in Figure 5 are as shown in Figure 6, and are seen to decrease as the temperature increases. If the calculation is made on a thermo-electric basis then the value is approximately in accordance with predictions even at low temperatures, whereas the values quoted from measurements of conductivity are too high at lower temperatures, but tend to the same value ($\simeq 5$) as the results from thermo-electric measurements at higher temperatures. This value is in fair agreement with the mass action value of 6, Baumbach and Wagner's (1934) conductivity value of 4, and the value of 4·8 calculated by the present author (Hogarth 1948 b) from the spot readings of

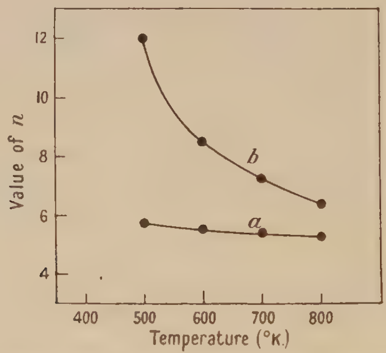


Figure 6. The variation of the constant n with temperature deduced from a, thermo-electric and b, conductivity measurements on green nickel oxide.

thermo-electric power given by Baumbach and Wagner (1934). The results are in accordance with the general theory for a deficit semiconductor as previously presented (Hogarth 1948 c).

In Figure 5 (*b*) the slope of the curves is seen to increase with pressure of oxygen, indicating an apparent increase in activation energy. If these energy values are plotted against log P_{O_2}, a straight line is obtained as shown in Figure 7.

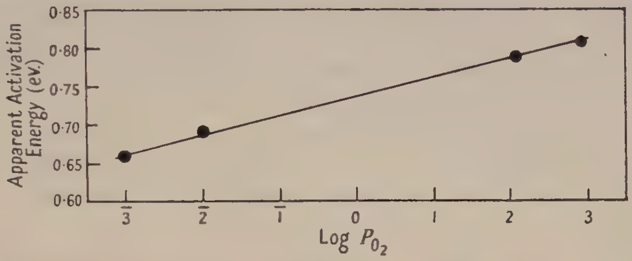


Figure 7. Apparent activation energy as a function of oxygen pressure for green nickel oxide.

One interpretation of this curve is that the slope at normal conditions of air pressure represents the usual thermal activation energy, whereas the other slopes provide a measure of the heat of solution or evolution of oxygen in NiO. The logarithmic variation is in accord with the theories of gaseous absorption in solids.

The results for the variation of thermo-electric power with oxygen pressure shown in Figures 1, 2 and 5 (*c*) indicate that the predictions of theory concerning the numerical values of the constant n are more nearly verified in practice for the specimens of higher purity. Also, the agreement in value between the slopes of the lines of Figures 2 and 5 (*c*) indicates that the results of two distinct measuring techniques are consistent with each other.

(ii) *Cadmium Oxide*

The results of variation with oxygen pressure of the thermo-electric power and electrical conductivity in the temperature range 200–550° c. are in agreement with previous investigations on the subject (Hogarth and Andrews 1949, Baumbach and Wagner 1933). Thus the general form of equations (2) and (3) may be considered verified, both sets of measurements having been made on the same specimen. Again the values of n calculated from thermo-electric measurements are higher than expected and the conductivity values ($\simeq 6$) are of the expected order of magnitude and in agreement with the results of Baumbach and Wagner. In so far as the conductivity increases and the thermo-electric power decreases as the oxygen pressure is reduced, CdO behaves as expected for an excess semiconductor in this temperature range. This is in agreement with the sign of the Hall coefficient (Hogarth 1951, Wright 1949), but the actual mechanism of conduction will be considered below.

In the equations given for the variation of thermo-electric power with temperature for CdO (Andrews 1947), $dE/dT = -a + b/T$ and for the empirical results of the present measurements $dE/dT = -(C_T \log P_{O_2} + \text{const.})$, where a, b and C_T are constants whose values have been calculated, the values found are of the same order of magnitude as those found previously (Hogarth and Andrews 1949), and vary in a similar manner. The consistence of these results again indicates the validity of both methods of measurement of thermo-electric power.

(iii) *Differential Thermo-electric Effects*

Wagner (1933) showed by a thermodynamic argument that the 'differential thermo-electric power' $\Delta(dE/dT)$ developed between two specimens of the same semiconductor maintained at different vapour pressures so that their conductivities σ_1 and σ_2 were different, was given by $\Delta(dE/dT) = \mp(\boldsymbol{k}/\boldsymbol{e})\log(\sigma_1/\sigma_2)$. With his collaborators, this relation was verified approximately for Cu_2O, ZnO and NiO at one or two temperatures. Croatto (1949) has verified the relation for CdO at 600° c. and finds that the calculated value is $5{\cdot}67 \times 10^{-5}$ v/° c. whereas the differential value (determined by the use of the law of intermediate metals) is $4{\cdot}8 \times 10^{-5}$ v/° c. for the pressures used in his experiments.

The relation has been examined using the experimental results of this paper. Table 1 shows the values for CdO, and agreement between the calculated and experimental values is good, particularly at the higher temperatures.

The corresponding results for NiO are shown in Table 2. Here agreement is less satisfactory although again it is better at the higher temperatures.

Table 1. Values for the Differential Thermo-electric Power in CdO calculated (*a*) directly, (*b*) from measurements of electrical conductivity, at 0·001 mm. air and 777 mm. oxygen. (Results from Figures 4 (*a*) and (*b*))

$10^4/T$	13	14	15	16	17
T (° K.)	769	714	667	625	588
(*a*) $\Delta(dE/dT)$ (v/°c. $\times 10^4$)	0·25	0·23	0·22	0·19	0·17
(*b*) $(\boldsymbol{k}/\boldsymbol{e})\ln(\sigma_1/\sigma_2)$ (v/°c. $\times 10^4$)	0·224	0·223	0·214	0·215	0·211

Table 2. Values for the Differential Thermo-electric Power in Green NiO calculated (*a*) directly, (*b*) from measurements of electrical conductivity at 0·005 mm. air and 767 mm. oxygen. (Results from Figures 5 (*a*) and (*b*))

T (° K.)	500	600	700	800
(*a*) $\Delta(dE/dT)$ (v/°c. $\times 10^4$)	2·05	2·11	2·15	2·24
(*b*) $(\boldsymbol{k}/\boldsymbol{e})\ln(\sigma_1/\sigma_2)$ (v/°c. $\times 10^4$)	1·45	1·72	1·82	2·05

(iv) *Conduction Mechanism for Cadmium Oxide*

No measurements on the conduction properties of single crystals of CdO have so far been reported and all results have been obtained with thin films or pressed powder specimens. CdO has a high electrical conductivity (1 to 10^3 ohm^{-1} cm^{-1} at room temperature), and although results for its temperature variation are few, it is generally believed that the conductivity decreases with increasing temperature much as for a metal. Conduction measurements on thin films (Bädeker 1907, Preston 1950) are in accordance with the above, and Wright (1949) has shown that the temperature variation of conductivity is 'metallic' and, in addition, that the Hall coefficient is constant from room temperature up to 500° c., indicating a constant concentration of free electrons. The values of $1{\cdot}7 \times 10^{19}$ cm^{-3} and 250 ohm^{-1} cm^{-1} for the concentration and conductivity respectively indicate a large deviation from stoichiometric proportions in his specimens, and the method of preparation adopted would have resulted in a considerable reduction of the CdO. Faivre (1944) has shown that a darkening of CdO is associated with a high proportion of excess Cd and that such specimens should be considered as *impurity metal*, much as described by Shifrin (1944).

On this basis, Wright has assumed a constant concentration of free electrons, and by evaluating certain of Shifrin's distribution functions given in integral form, has been able to produce equations which fit exactly his experimental results.

Measurements of Hall effect on non-reduced specimens made by the present author (Hogarth 1951) have indicated a constant concentration of free electrons ($\simeq 2 \times 10^{18}\,\mathrm{cm}^{-3}$) up to about 150° c. and then a sharp increase as the temperature is further raised. This result together with that of Figure 3 would indicate that at some temperature (determined by the relative impurity content of the specimen concerned) there is a conductivity minimum for CdO. Previously quoted results were probably obtained by measurements at temperatures below that at which the minimum would have occurred. Measurements of conductivity against temperature on CdO specimens used in Hall effect studies have indicated that the mobility decreases slowly with temperature, so that the increase in conductivity and decrease in Hall coefficient which occur at about 150° c. is due to an increase in the free carrier concentration.

Meyer (1944) has discussed the band model for CdO and has classified it as an *overlapping* semiconductor in which the upper filled band and next higher empty band overlap and result in a metallic type of conduction. On the basis of the results described above it would seem more probable that the band diagram should be as follows. The excess Cd in the form of Cd^{++} ions and free electrons is present in sufficient quantity to form a narrow band about 0·4 ev. below the bottom of the conduction band which will be practically empty at normal temperatures. Metallic conduction is able to take place in this impurity band much as in Meyer's overlapping bands which leads to an incompletely filled impurity band, or by means of transitions from one non-localized impurity state to another as discussed by Hung (1950) for germanium at low temperatures. At higher temperatures (above 150° c. in this case) it is possible for electrons to make the transition from this impurity band to the conduction band, requiring an activation energy of about 0·4 ev. to do so. Such a scheme can be fitted into an extension of Shifrin's theory as indicated by Busch and Labhart (1946) for SiC. The type of conduction behaviour described above is observed for PbS (Dunaev 1947) and a similar mechanism might be postulated in this case.

ACKNOWLEDGMENTS

Most of the experimental work described in this paper was performed at Queen Mary College, London, under the supervision of Dr. J. P. Andrews and with encouragement from Professor H. R. Robinson. For recent stimulating discussions I should like to express my thanks to Dr. R. W. H. Wright and Dr. H. K. Henisch. A maintenance grant from the Department of Scientific and Industrial Research during the period 1946–1948 is acknowledged with appreciation.

REFERENCES

Anderson, J. S., and Morton, M. C., 1945, *Proc. Roy. Soc.* A, **184,** 82.

Andrews, J. P., 1947, *Proc. Phys. Soc.*, **59,** 990.

Bädeker, K., 1907, *Ann. Phys., Lpz.*, **22,** 749; 1909, *Ibid.*, **29,** 566; 1911, *Phys. Z.*, **13,** 1080.

Baumbach, H. H. von, and Wagner, C., 1933, *Z. Phys. Chem.* B, **22,** 199; 1934, *Ibid.*, **24,** 59.

Busch, G., and Labhart, H., 1946, *Helv. phys. Acta*, **19,** 463.

Croatto, U., 1949, *Ric. Sci.*, **19,** 1324.

Dunaev, J., 1947, *C.R. Acad. Sci., U.R.S.S.*, **21,** 55.
Faivre, R., 1944, *Ann. Chimie,* **19,** 58.
Gudden, B., 1934, *Ergeb. Exakt. Naturwiss.*, **13,** 223.
Henisch, H. K., 1949, *Thesis* (Reading).
Hogarth, C. A., 1948 a, *Phil. Mag.*, **39,** 260; 1948 b, *Nature, Lond.*, **161,** 60; 1948 c, *Thesis* (London); 1951, *Nature, Lond.*, **167,** 521.
Hogarth, C. A., and Andrews, J. P., 1949, *Phil. Mag.*, **40,** 273.
Hung, C. S., 1950, *Phys. Rev.*, **79,** 727.
Meyer, W., 1944, *Z. Elektrochem.*, **50,** 274.
Nagel, K., and Wagner, C., 1934, *Z. phys. Chem.* B, **25,** 71.
Prasad, M., and Tendulkar, M. G., 1931, *J. Chem. Soc.*, **1,** 1403.
Preston, J. S., 1950, *Proc. Roy. Soc.* A, **202,** 429.
Shifrin, K., 1944, *J. Phys. U.S.S.R.*, **8,** 242.
Sproull, R., and Tyler, W., 1951, *Proceedings of an International Conference on the Properties of Semiconducting Materials* (London: Butterworth's Scientific Publications).
Wagner, C., 1933, *Z. phys. Chem.* B, **22,** 195.
Wright, R. W. H., 1949, *Thesis* (London).

The Conductivity–Temperature Characteristics of Lead Sulphide: The Influence of Oxygen and of the Rate of Heating

By W. EHRENBERG and J. HIRSCH *
Birkbeck College, University of London

MS. received 5*th December* 1950, *and in amended form* 29*th January* 1951

ABSTRACT. In general, the conductivity–temperature curves of impurity semiconductors should be affected by the rate of heating. An arrangement is described by which such curves can be recorded with rates of heating up to 5,000 centigrade degrees per second. Extensive tests made on lead sulphide films in an oxygen atmosphere lead to conclusions about the stages of oxidation.

§ 1. INTRODUCTION

When a semiconductor is heated, the distribution of electrons between free and bound states changes in favour of the free states. This process is assumed to be instantaneous. In addition, atomic processes, such as the displacement of the constituents of the lattice and their exchanges between lattice and ambient atmosphere, modify the number and nature of the bound states; very little is known about the rates of the atomic processes. The changes in resistance of a semiconductor with temperature will, therefore, in general be due to a combination of both processes, so that the resistance–temperature characteristics should depend on the rate of heating : if the temperature changes sufficiently rapidly, electronic processes alone can take place; if the temperature changes sufficiently slowly, the semiconductor will always be in equilibrium. Only in these limiting cases will the characteristics be entirely reversible. If two different rates of temperature change lead to the same characteristic, both rates represent the same limiting case; if they lead to different characteristics, the slower change at least is accompanied by an atomic process.

Hence it appeared worth while to develop a technique for measuring conductivity at widely different rates of heating and to apply it to an impurity semiconductor such as lead sulphide.

* Now at Purdue University, Lafayette, Indiana. This paper is a summary of J. Hirsch's thesis for the Ph.D. degree approved by the University of London (June 1950).

§ 2. EXPERIMENTAL ARRANGEMENT

In order to obtain a high rate of heating, the semiconductor was deposited as a narrow strip (Figure 1) on one side of a thin mica lamella next to a platinum strip (Pt thermometer) whose resistance as a function of temperature was known; the back of the mica lamella (Figure 1 (*a*)) was platinized (heater). On account of the small heat capacity a current of a few amperes produced a rate of rise of temperature of up to 5,000 centigrade degrees per second. Hence the range between 100° K. and 600° K. could be covered in slightly more than 0·1 second. An elaborate circuit enabled the heater current to be switched on for the duration of one sweep of the time base of a cathode-ray tube; simultaneously the brightness of the trace of the cathode-ray tube was increased. The voltages developed across the platinum thermometer and the specimen by small constant currents (measuring temperature and resistance of the specimen) were, after amplification, fed in rapid alternation to the vertical deflection plates of the cathode-ray oscillograph so that, on pressing a button, the whole information required for a resistance–temperature curve could be recorded photographically. Typical records are shown in Figure 3. Altogether 420 records were made during the investigation. Alternatively a small steady current through the heater could maintain any desired temperature.

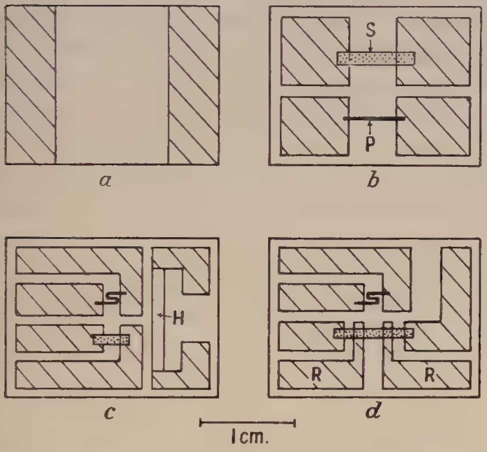


Figure 1. Specimen carrier. Mica 0·001 in. thick. Three types have been used. The back *a* carries in all cases the main heater, a platinum film sputtered across gold ‘bus bars’ deposited by the ‘liquid bright gold’ process. The first type of front is shown in *b*; it has only two pairs of electrodes, one connected by the specimen under tests S and the other by a narrow platinum strip P. *c* shows a specimen permitting the measurement of a thermo-E.M.F. The additional heater H provides a temperature gradient across the PbS film. *d* shows another six-electrode arrangement with which the resistivity of the specimen can be measured by probes R.

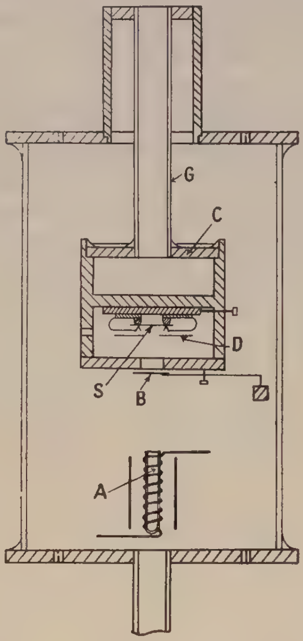


Figure 2. Specimen chamber. The copper box C is suspended from the top by a german silver tube G, through which its upper half can be filled with liquid air. In its lower compartment the specimen carrier S is clipped by contact springs. PbS is evaporated from the quartz crucible A through the aperture B in the box and the stencil D on to the mica blank.

The specimens were prepared by inserting the mica lamella, complete but for the semiconductor, in the specimen chamber (Figure 2). It is seen that the specimen was placed in a copper box which could be cooled with liquid air. After the chamber was evacuated the lead sulphide was evaporated from a silica crucible on to the lamella through a stencil. During evaporation the resistance of the deposit and its temperature could be measured. In the course of the subsequent tests the specimens were annealed at various temperatures, first *in vacuo*, then at increasing pressures of oxygen, for periods of time varying from a few minutes to several weeks. At frequent intervals the dependence of resistance on temperature, and the effect of varying the rate of heating, were studied, and on some specimens (see Figure 1 (*c*)) the thermo-electric power was also observed. It was invariably found that measurements with probes (Figure 1 (*d*)) and direct measurements (Figure 1(*b*)) led to the same resistance. Lead sulphide of different grades of purity was used, but the purity of the compound did not appear to affect the results. This is not astonishing, as sublimation is known to be a very effective means of removing impurities. On the other hand, sublimation must be expected to cause a certain amount of dissociation which would result in a Pb/S ratio greater than one in the layer (Jellinek and Deubel 1929, Schenk and Albers 1919).

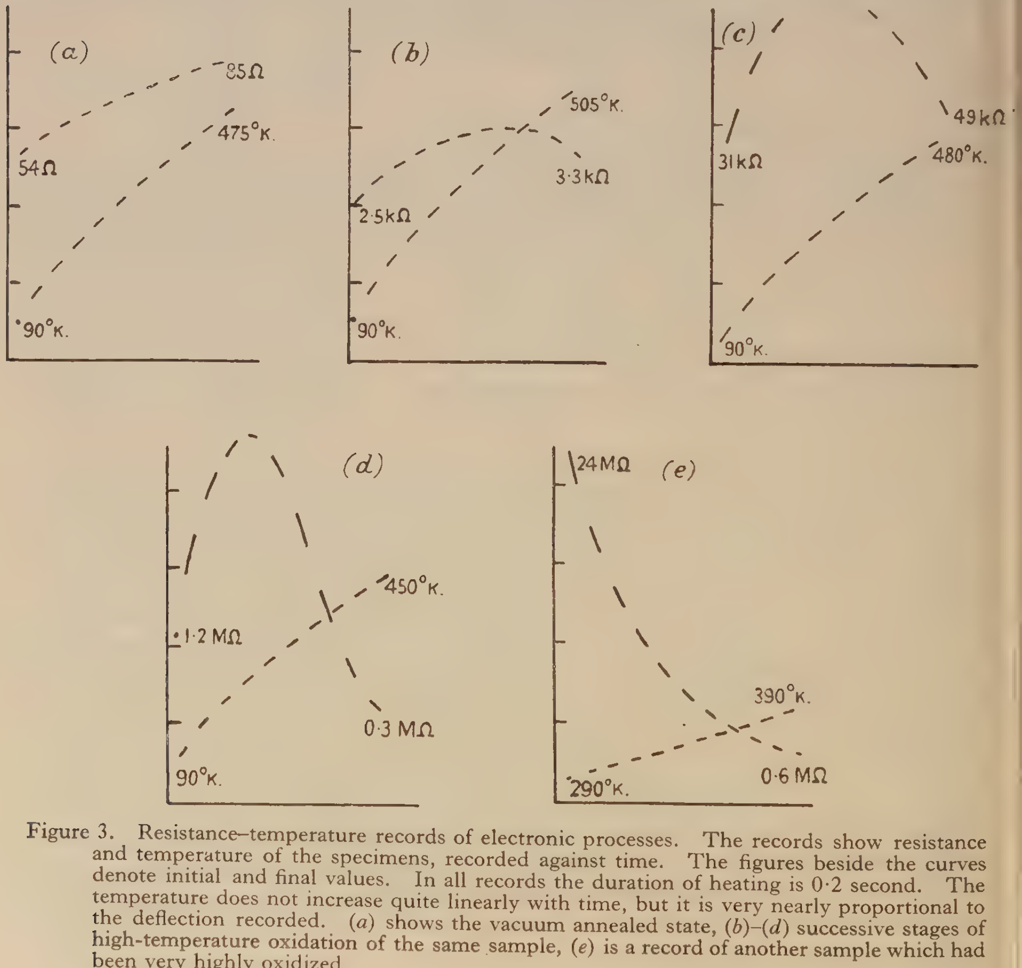


Figure 3. Resistance–temperature records of electronic processes. The records show resistance and temperature of the specimens, recorded against time. The figures beside the curves denote initial and final values. In all records the duration of heating is 0·2 second. The temperature does not increase quite linearly with time, but it is very nearly proportional to the deflection recorded. (*a*) shows the vacuum annealed state, (*b*)–(*d*) successive stages of high-temperature oxidation of the same sample, (*e*) is a record of another sample which had been very highly oxidized.

For the final tests, however, PbS of highest purity was used. This contained as only significant impurity either less than 0·001% copper (batch I) or less than 0·001% chloride plus less than 10^{-5} metal (batch II). After completion of the tests the thickness of the layer was determined colorimetrically by dissolving a known area of the deposit and re-precipitating it as colloidal lead sulphide.

§ 3. RESULTS

Though the results were not completely reproducible, specimens which were evaporated within a wide range of conditions behaved very much alike. This standard behaviour resulted when the mica blank was kept above room temperature during evaporation, and the evaporation was rapid and carried on until the films were at least 0·3 μ thick, i.e. when the condition of evaporation favoured clean deposits and large grains (Wilman 1948). The results given refer to films of standard behaviour.

(i) *Oxygen-Free Lead Sulphide*

A freshly evaporated film is in an unstable state, but extended annealing *in vacuo* at temperatures up to 500° K. brings it into a very stable condition with a low-temperature resistance about half that immediately after evaporation. As long as the specimen is kept *in vacuo*, identical, reproducible, and hence purely electronic, resistance–temperature characteristics are observed at all rates of change of temperature. A typical record is shown in Figure 3 (*a*). The resistivity is low and increases linearly with temperature; also, the specimens have a negative thermo-electric power corresponding to conduction by electrons. At room temperature one specimen had a conductivity of 800 $\text{ohm}^{-1}\,\text{cm}^{-1}$, but a value of 100 $\text{ohm}^{-1}\,\text{cm}^{-1}$ was more usual. The latter value is about equal to values found by Hintenberger (1942) for films deposited on glass; the former is near to that given by Eisenmann (1940) for large specimens obtained from the melt. This indicates that no systematic thin film effects need be assumed.

From expressions given by Shifrin (1944) and Ehrenberg (1950) for the resistivity of simple non-ionic semiconductors, resistivity–temperature curves have been constructed, with parameters N, the number of impurity centres, and $\Delta\epsilon$, the energy difference between the bottom of the conduction band and the impurity level. A comparison of these curves with the experimental ones indicates values of $\Delta\epsilon \leqslant 0{\cdot}01$ ev. and $N \sim 10^{20}\,\text{cm}^{-3}$. The impurity is presumably excess lead, i.e. S^{--} vacancies or F-centres. The observed thermo-electric power is only one-sixth of that predicted by Shifrin's (1944) expression. After exposure to oxygen at a temperature not exceeding 400° K., the condition described can be restored by renewed annealing *in vacuo.*

(ii) *Oxidized Lead Sulphide*

When a specimen is oxidized at a given pressure and temperature, its resistivity at this temperature rises rapidly at first and approaches asymptotically a limiting value. The resistance–temperature characteristics can then be measured at the same pressure at a medium rate (within minutes) while the specimen cools, and rapidly (within a fraction of a second) by heating the specimen again. The two curves are identical and reproducible provided the temperature of oxidation is not exceeded on reheating: the measurements thus disclose purely electronic changes corresponding to the state of oxidation at the end of the heat treatment. The effects of a very slow rate of cooling (extending over hours) will be discussed below under the heading 'Mixed processes'.

We found that the effects of oxidation at or above 400° K. (high-temperature oxidation) had to be distinguished from those of low-temperature oxidation (below 400° K.).

Electronic processes after high-temperature oxidation.

In Figure 3 (*b*)–(*e*) is reproduced a series of records obtained after extended oxidation at high temperature and increasing pressures. It is seen that the resistivity–temperature characteristic ceases to be monotonic; a maximum of resistivity occurs well below the temperature of oxidation, which, with heat treatment at higher pressures, becomes more marked and shifts to a lower temperature. At low temperatures the shallow F-centres not yet destroyed by oxidation still give rise to quasi-metallic conductivity. Beyond the maximum the contribution from substituted oxygen centres predominates and the conductivity is proportional to $\exp(-\Delta\epsilon/2\boldsymbol{k}T)$ with $\Delta\epsilon \sim 0{\cdot}3$ ev. As the oxidation of the specimens progresses by annealing at increasing pressures, and thus the replacement of F-centres by O^{--} ions becomes more complete, the conductivity decreases steadily to about 10^{-3} ohm^{-1} cm^{-1} and 10^{-5} ohm^{-1} cm^{-1}, measured at 290° K. and 90° K. respectively, i.e. by a factor 10^{-5} to 10^{-6}. The large decrease of conductivity with oxidation at high temperature has been observed and interpreted in this way by Hintenberger (1946) and Anderson (1948).

Electronic processes after low-temperature oxidation.

Exposure to oxygen of a vacuum-annealed specimen at temperatures below 400° K. leads to a decrease of conductivity by a factor of only 10^{-1} to 10^{-2} and to resistance–temperature curves which are qualitatively different from the ones described above: the linear resistance–temperature relation, characteristic of vacuum-annealed samples, is retained. The drop in conductivity becomes greater the higher the pressure and the lower the temperature of exposure. This behaviour should be associated with adsorption of oxygen at the surface of the grains to which diffusion can take place along grain boundaries even at low temperatures. Such adsorption has already been observed by Anderson and Morton (1945) on stannous sulphide, and by Anderson (1948) on lead sulphide, but it is not known in detail how the adsorption of oxygen affects the conductivity. Anderson (1948) has discussed this adsorption, but in his interpretation he neglects the effect of space charge necessarily associated with his model.

Mixed processes.

At temperatures higher than the temperature of annealing it is not possible to observe purely electronic resistance–temperature characteristics, even at the most rapid rate of heating available. An example of this failure is shown in Figure 4 (*a*). The specimen had been exposed to oxygen (10^{-2} mm. Hg for 20 hours) at room temperature, and during this exposure its resistance had increased about fifty-fold. It was then subjected to pulsed heating at the same oxygen pressure. The record shows the still quasi-metallic resistance–temperature characteristics between 90° K. and room temperature. As the temperature during the pulse increases above room temperature, oxygen is desorbed with increasing speed and the resistance decreases. The resemblance between Figures 4(*a*) and 3(*d*) is superficial; at the end of the experiment (Figure 4(*a*)) the resistance at 90° K. was found to have fallen to about one-tenth of its initial value. Figure 4(*b*) represents the variation of resistance with temperature of another specimen which had been annealed *in vacuo*; the figure records the effect of rapid heating (0·2 second) immediately after admission of oxygen. One notices an alternation between concave curvature, i.e. intake of oxygen, and convex curvature, i.e. desorption of oxygen. Similar very fast processes have been observed by Anderson.

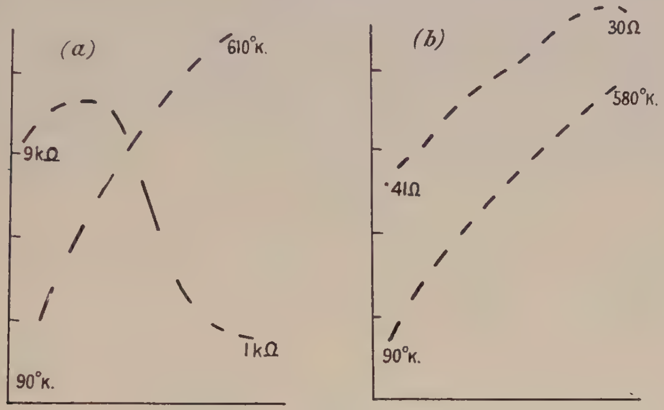


Figure 4. Resistance–temperature records. Mixed processes.

On the other hand, very slow atomic processes take place in specimens exposed to oxygen at the pressure, but below the temperature, of annealing. They are, as stated above, not observed within minutes, but may extend over hours, days and even weeks. It has been noted that the room-temperature conductivity appertaining to a state reached after high-temperature oxidation is correctly measured after rapid cooling, and that the conductivity steadily decreases with increasing oxidation. This is contrary to Hintenberger's (1946) statement that by high-temperature oxidation at pressures above 1 mm. Hg the conductivity increases again. An increase is in fact observed only if the specimens are left exposed to oxygen below 400° K. This increase is accompanied first by a decrease in magnitude of the thermo-electric power (which assumes a large negative value as the result of the high-temperature oxidation), then by a reversal of its sign and a growth in the magnitude of the now positive thermo-electric power. The effect again increases with increasing pressure and with decreasing temperature; this state of increased conductivity is very unstable thermally: the effect of several days' oxidation at room temperature and atmospheric pressure of oxygen could be undone by a single heating pulse with a ceiling temperature of about 400° K. Moreover, we found no other way of inducing positive hole conductivity in our specimens.

We conclude that Hintenberger allowed too long a time between high-temperature oxidation and measurement, and that oxidation at high temperatures leads only to a substitution of oxygen for the S^{--} vacancies (F-centres), with a resulting decrease in electronic conductivity. The increase in conductivity is then interpreted as due to a release of positive holes from oxygen adsorbed at grain boundaries.

ACKNOWLEDGMENTS

The work described was carried out at the Research Laboratories of Birkbeck College, London. The authors wish to thank Professor J. D. Bernal for his encouragement and the Nuffield Foundation for generous grants. They also wish to express their thanks to Dr. R. W. Pittman of the Chemistry Department of Birkbeck College and to Messrs. Levy and West Ltd. for the preparation of the pure lead sulphide referred to as batches I and II. One of the authors (J. H.) was in receipt of a grant from the Department of Scientific and Industrial Research for a part of the time during which the work was carried out.

REFERENCES

Anderson, J. S., 1948, *Discussions of the Faraday Society*, **4**, 163.
Anderson, J. S., and Morton, M. C., 1945, *Proc. Roy. Soc.* A, **184**, 83.
Ehrenberg, W., 1950, *Proc. Phys. Soc.* A, **63**, 75.
Eisenmann, L., 1940, *Ann. Phys., Lpz.* (5), **38**, 121.
Hintenberger, H., 1942, *Z. Phys.*, **119**, 1; 1946, *Z. Naturforsch.*, **1**, 13.
Jellinek, K., and Deubel, A., 1929, *Z. Elektrochem.*, **35**, 451.
Schenk, R., and Albers, A., 1919, *Z. Allg. Anorg. Chemie*, **105**, 164.
Shifrin, K., 1944, *J. Phys. U.S.S.R.*, **8**, 242.
Wilman, H., 1948, *Proc. Phys. Soc.*, **60**, 117.

Piezoelectric Behaviour of Partially Plated Square Plates Vibrating in Contour Modes

By R. BECHMANN and P. L. PARSONS
Post Office Research Station, Dollis Hill, London

MS. received 8*th November* 1950

ABSTRACT. In extension of a previous paper, the piezoelectric behaviour of partially plated square plates vibrating in contour modes is used to check the theoretical solution for the motion. For the longitudinal mode 2 recently described by Ekstein a particularly sensitive test of the theory is possible. A critical electrode arrangement should cause the piezoelectric excitation to vanish. This is confirmed by experiment.

§ 1. INTRODUCTION

As already mentioned in an earlier paper (Bechmann 1951), the theoretical solution for the motion of a plate vibrating in a contour mode can be verified from measurements of the equivalent electric circuit. The present paper shows that this verification can be made more completely by measurements on plates whose electrodes extend over only part of the faces.

The dynamic capacitance in the equivalent electric circuit can be calculated for any shape of electrodes from the solution of the motion. Hence by measurements of the dynamic capacitance, using different electrode arrangements, the theoretical solution for the motion can be tested. This effect is discussed below for the solution of the three longitudinal modes given by Ekstein (1944) and for the shear mode solution given by Mähly and Trösch (1947), with experimental results for longitudinal modes 1 and 2. The notation used is taken from the preceding paper, which contains a summary of these solutions in Tables 1 and 2.

The effects in partially plated crystals may also be of interest in practical applications of piezoelectric elements, such as in channel filter circuits. For instance, a large ratio of dynamic capacitance to parallel capacitance is required in filter design. That a suitable choice of electrode arrangement may increase this ratio above the value obtained with a fully plated crystal has previously been shown for the case of longitudinally vibrating bars (Kamayachi and Ishikawa 1940, Cady 1946) and shear vibrating square plates (Bechmann 1943). The variation of electrode size may also be used to control the dynamic capacitance of a filter crystal in place of the normal method of adjusting the thickness. This is of special interest when it is necessary to avoid a particular thickness–width ratio which gives undesirable coupling with other modes.

§ 2. CALCULATION OF EQUIVALENT ELECTRIC CIRCUIT OF PARTIALLY PLATED SQUARE PLATES

The calculation of the equivalent electric circuit for the various modes of square plates is outlined in the paper mentioned above (Bechmann 1951). The dynamic capacitance C_k follows from the equivalence of the electric and elastic energy:

$$C_k = I^2_p / 2\omega^2 T \qquad \ldots\ldots(1)$$

where I_p is the peak value of the total current due to the piezoelectric polarization over the area of the plate covered by electrodes. This value is obtained by integrating the current density over the electrode area (equation (15), Bechmann 1951), using the solution for the motion to evaluate the strain. In the following it is assumed that the elastic energy T of the vibrating crystal element and hence the angular frequency ω remain unchanged when the size of the electrodes is reduced. This assumption is an approximation which leads to error for very small electrodes. The results are also modified slightly by fringing effects at the edges of the plated area.

It is convenient to introduce the capacitance constant Γ defined by $\Gamma = C_k t/a$ where t is the thickness of the plate, and a the area of electrode. Since the parallel capacitance in the equivalent electric circuit is $\epsilon a/4\pi t$ where ϵ is the dielectric constant, the ratio of dynamic to parallel capacitance is $4\pi\Gamma/\epsilon$. The capacitance constant Γ is given by

$$\Gamma = \rho\beta d^2 N^2 \qquad \ldots\ldots(2)$$

(equation (17), Bechmann 1951), where ρ is the density, d the effective piezoelectric coefficient and N the frequency constant. β is a numerical factor which depends on the mode of vibration and the arrangement of the electrodes. In the case of fully plated crystals, values of β for the three longitudinal and the shear mode of square plates have been given in the previous paper.

The equivalent electric circuit of a plate with reduced electrode size is now discussed in terms of the ratio β/β_0, where the suffix zero refers to the value for the fully plated crystal. From equations (1) and (2) it follows that

$$\frac{\beta}{\beta_0} = \frac{\Gamma}{\Gamma_0} = \frac{C_k}{C_{k_0}}\frac{a_0}{a} = \frac{I_p{}^2}{I_{p_0}{}^2}\frac{a_0}{a}. \qquad \ldots\ldots(3)$$

As the size of the electrode tends to zero, Γ/Γ_0 and β/β_0 must fall to zero, since I_p is the surface integral of a finite quantity over the area a.

In the following calculations, the origin of coordinates is taken at the centre of the square plate of side l_0. The x and y axes are in the plane of the plate and parallel to the edges (Figure 1).

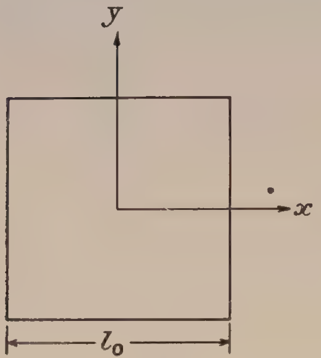


Figure 1. Coordinate system for square plate.

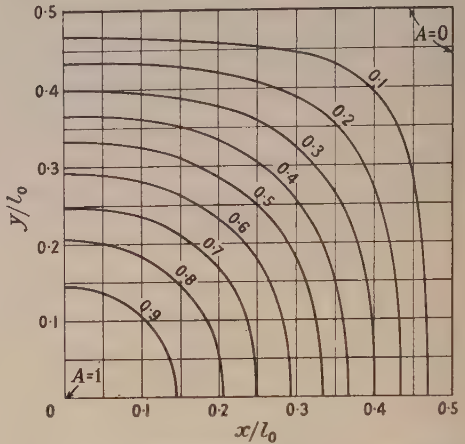


Figure 2. Longitudinal mode 1 of square plate. One quadrant of cos kx cos $ky=A$. Values for A are shown against the curves.

§ 3. LONGITUDINAL MODE 1 OF SQUARE PLATES

From the solution for the longitudinal mode 1, the piezoelectric polarization at any point (x,y) is proportional to $\cos kx \cos ky$, where $k=\pi/l_0$. The curves

$$\cos kx \cos ky = A, \qquad \ldots\ldots(4)$$

where A is a quantity proportional to the polarization, join points making equal contributions to the electric current (Figure 2). $A=0$ represents the crystal boundary and $A=1$ the centre. The total current is proportional to

$$\int\int_a \cos kx \cos ky \, dx \, dy \qquad \ldots\ldots(5)$$

the integral being taken over the area of the electrode.

For an electrode forming a symmetrical strip of width l the limits of the integral become: $x=\pm\frac{1}{2}l$, $y=\pm\frac{1}{2}l_0$ and hence from equation (3)

$$\beta/\beta_0=[\sin^2(\pi l/2l_0)]/(l/l_0). \qquad \ldots\ldots(6)$$

The function β/β_0 reaches a maximum value of 1·138 at $l/l_0=0{\cdot}742$. The expression (6) is identical with that for a longitudinally vibrating narrow bar of length l_0 carrying an electrode of length l (Cady 1946, p. 303).

For a central square electrode of side l, the limits of the integration are $x=\pm\frac{1}{2}l$, $y=\pm\frac{1}{2}l$ and

$$\beta/\beta_0=[\sin^4(\pi l/2l_0)]/(l/l_0)^2. \qquad \ldots\ldots(7)$$

The maximum value again occurs at $l/l_0=0{\cdot}742$, where β/β_0 now rises to 1·296.

For a central circular electrode of radius $\rho \leqslant \frac{1}{2}l_0$, from the calculation given in the Appendix $\beta/\beta_0=(\frac{1}{2}\pi)^3[J_1(\frac{1}{2}\pi\sqrt{2}\rho/\rho_0)]^2$ where $J_1(x)$ is the Bessel function of the first order. At $\rho/\rho_0=0{\cdot}828$ a maximum value $\beta/\beta_0=1{\cdot}312$ is reached.

The calculated values of β/β_0 for these three cases are shown in Figure 3 as a function of l/l_0 or ρ/ρ_0. In the special case of the circular electrode with diameter equal to the side of the square ($\rho/\rho_0=1$), $\beta/\beta_0=1{\cdot}184$, the static capacitance is reduced by a factor $\pi/4$ and the dynamic capacitance C_k is $0{\cdot}930\ C_{k0}$. The maximum possible value of β/β_0 for this particular mode is obtained by plating an area bounded by a member of the curves (4). This value is $\beta/\beta_0 = 1{\cdot}33$, obtained for $A = 0{\cdot}30$, and is only slightly greater than the maximum obtainable using circular electrodes.

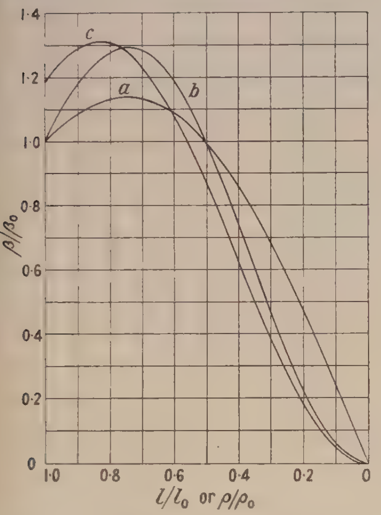


Figure 3. Variation of β/β_0 with electrode size for square plate, longitudinal mode 1, electrode forming : (*a*) symmetrical strip, (*b*) central square, (*c*) central circle.

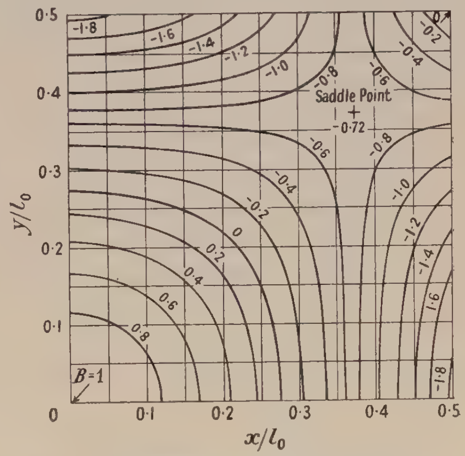


Figure 4. Longitudinal mode 2 of square plate. One quadrant of $[4/(4-\pi)]\cos kx \cos ky - [\pi/2(4-\pi)] \times (\cos kx + \cos ky) = B$. Values of B are shown against the curves.

§ 4. LONGITUDINAL MODE 2 OF SQUARE PLATES

For longitudinal mode 2 the piezoelectric polarization at any point (x, y) is proportional to $\cos kx \cos ky - 8(\cos kx + \cos ky)/\pi$. The curves

$$[4/(4-\pi)]\cos kx \cos ky - [\pi/2(4-\pi)](\cos kx + \cos ky) = B$$

(Figure 4) join points making equal contribution to the electric current. B is a quantity proportional to the piezoelectric polarization, adjusted to the value unity at the centre of the plate. The expression for B consists of two terms of opposite sign and this gives a characteristic effect for this mode: B is positive over a region surrounding the centre but is negative over the greater part of the plate. When the crystal is fully plated, the contribution of the outer region to the current outweighs that of the inner. In consequence, reduction of the electrode to a particular size causes the excitation of the plate to vanish. If the electrode size is further reduced, the excitation reappears. This compensating effect depends critically on the electrode size and lends itself to a sensitive test of the distribution on which the calculation is based.

For the case of a central square symmetrical electrode of side l, β/β_0 becomes

$$\frac{\beta}{\beta_0} = \frac{[\sin^2 \pi l/2l_0 - \pi^2 l/8l_0 \sin \pi l/2l_0]^2}{(\frac{1}{8}\pi^2 - 1)^2 (l/l_0)^2}. \qquad \ldots\ldots(8)$$

β/β_0 is zero for $l/l_0 = 0{\cdot}748$, and rises again to a maximum of $\beta/\beta_0 = 0{\cdot}351$ at $l/l_0 = 0{\cdot}400$.

For a central circular electrode of radius $\rho \leqslant \frac{1}{2} l_0$

$$\frac{\beta}{\beta_0} = \left(\frac{\pi}{2}\right)^3 \frac{[J_1(\pi\sqrt{2}\rho/2\rho_0) - (\pi\sqrt{2}/4)J_1(\pi\rho/2\rho_0)]^2}{(\frac{1}{8}\pi^2 - 1)^2} \qquad \ldots\ldots(9)$$

(for calculation see Appendix). The zero point for β/β_0 occurs at $\rho/\rho_0 = 0{\cdot}835$ and a maximum of $\beta/\beta_0 = 0{\cdot}362$ at $\rho/\rho_0 = 0{\cdot}459$. Figure 5 shows the function β/β_0 calculated for these two cases plotted against l/l_0 or ρ/ρ_0.

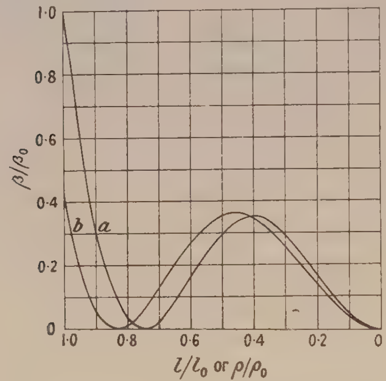


Figure 5. Variation of β/β_0 with electrode size for square plate, longitudinal mode 2, electrode forming :
(*a*) central square, (*b*) central circle.

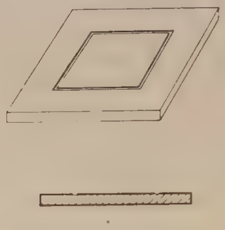


Figure 6. Electrode arrangement to reduce stray field.

In the special case of the largest possible circular electrode ($\rho/\rho_0 = 1$), $\beta/\beta_0 = 0{\cdot}419$; the static capacitance is reduced by a factor $\pi/4$ and the dynamic capacitance C_k is $0{\cdot}329\, C_{k0}$.

§ 5. MEASUREMENTS ON LONGITUDINAL MODES 1 AND 2

Measurements of frequency and dynamic capacitance have been made to test the calculated behaviour for modes 1 and 2, using square plates of ammonium dihydrogen phosphate (ADP) $Z_{0^\circ, 45^\circ}$ which can be excited in both modes 1 and 2. Preliminary measurements were made with various size square electrodes of evaporated silver, the outer part of the plate remaining unsilvered. The results indicated that the stray field at the edges of the silvered area modified the capacitance considerably, particularly in the case of mode 2. To avoid this effect, the electrode arrangement shown in Figure 6 was adopted. One face of the crystal plate carried a central silvered square which formed one electrode. This was separated by a narrow gap from silvering extending over the remainder of this face, the edges and the whole of the other face, to form the second electrode. The outer part of the crystal was thus shielded and the stray field reduced.

Results for modes 1 and 2 are shown in Figure 7 which gives the experimental values of C_k/C_{k0} plotted against l/l_0 for several electrode sizes. The theoretical curves, calculated from equations (3), (7) and (8) are included for comparison. For both modes 1 and 2 satisfactory agreement between measurement and calculation is found. As the solution for mode 1 has been previously verified from elastic and piezoelectric constants, this agreement indicates the reliability of the measurements. The results for mode 2, particularly the condition of zero excitation, confirm the distribution formula for mode 2 by Ekstein (1944).

The measured change of frequency due to partial plating was less than 0·3% in the case of mode 1 and less than 0·5% for mode 2. The mounting and plating processes cause an uncertainty estimated at 0·2% of the frequency, so the effect due to partial plating is small. To a first approximation the assumption of constant frequency made in §2 is thus justified, so that a more detailed treatment taking this change into account appears unnecessary.

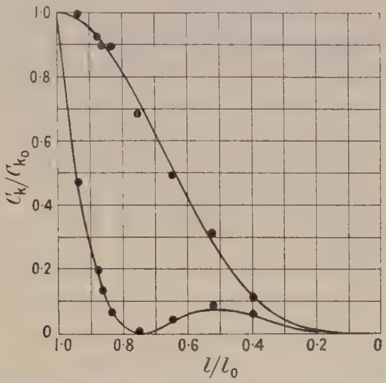


Figure 7. Measured and calculated variation of C_K/C_{K0} with electrode size, square electrodes, modes 1 and 2.

§6. LONGITUDINAL MODE 3 OF SQUARE PLATES

For the longitudinal mode 3 the piezoelectric polarization at any point (x, y) is proportional to $\cos kx + \cos ky$. In the case of the central square electrode, side l, the integration of this quantity leads to $\beta/\beta_0 = \sin^2(\pi l/2l_0)$. This function has no turning point. Since β/β_0 never exceeds unity, the capacitance ratio can not be increased above the value for the fully plated crystal.

The expression for the central circular electrode is $\beta/\beta_0 = \pi[J_1(\pi\rho/2\rho_0)]^2$ (see Appendix). The greatest value for β/β_0 is 1·009 at $\rho/\rho_0 = 1$.

Mähly (1945) has mentioned a slightly different solution for this mode, which depends on Poisson's ratio for the material, but the expression for the distribution function is not yet published. This would lead to a modified expression for β/β_0 which could be tested by experiment.

§7. CONTOUR SHEAR MODE OF SQUARE PLATES

The solution for the contour shear mode of square plates given by Mähly and Trösch (1947) depends on a material constant $\mu = s_l/s_s$, where s_l and s_s are respectively the longitudinal and shear elastic coefficients in the plane of the plate. This factor occurs in a correction term which is only important near the edges of the plate. The result for β_0, obtained by integrating the expression for the piezoelectric polarization over the whole plate, is independent of μ. However, if the integral is taken over only part of the plate, as required for partial plating, the result depends on this material constant. To check this theory by the effects of partial plating, measurements must therefore be made on a number of materials with different values of μ.

If the correction term is neglected, the piezoelectric polarization at any point (x, y) is proportional to $\cos k'x + \cos k'y$ where $k' = 2\kappa/l_0$, $\kappa = 2{\cdot}0288$ (Bechmann 1951, Table 1, equations (1) and (2)). This expression for the polarization differs from that for longitudinal mode 3 only in the value of k'.

In the case of the central square electrode, side l, this gives $\beta/\beta_0=[\sin^2(\kappa l/l_0)]/\sin^2\kappa$. This function has a maximum value $\beta/\beta_0=1{\cdot}243$ for $l/l_0=0{\cdot}774$. For the central circular electrode $\beta/\beta_0=\pi[\{J_1(\kappa\rho/\rho_0)\}/\sin\kappa]^2$. The maximum value is $\beta/\beta_0=1{\cdot}322$ for $\rho/\rho_0=0{\cdot}907$. If $\rho/\rho_0=1$ the dynamic capacitance C_{k} is increased to $1{\cdot}013\ C_{\mathrm{k}0}$, and the static capacitance is reduced by a factor $\pi/4$, $\beta/\beta_0=1{\cdot}290$.

ACKNOWLEDGMENT

Acknowledgment is made to the Engineer-in-Chief of the General Post Office for permission to make use of the information contained in this paper.

APPENDIX

EVALUATION OF CURRENT INTEGRAL FOR CIRCULAR ELECTRODES

The expression for the current integral for the various modes contain one or both of the following forms

$$\int_a\!\!\int \cos kx\cos ky\,dx\,dy\quad ..(\mathrm{A}\,1a);\qquad \int_a\!\!\int (\cos kx+\cos ky)\,dx\,dy\quad ..(\mathrm{A}\,1b)$$

where for circular electrodes a refers to the central circle of radius ρ. Introducing polar coordinates $x=r\cos\theta$, $y=r\sin\theta$, equation (A 1a) becomes

$$2\int_{r=0}^{\varrho}\int_{\theta=0}^{\pi}\cos(kr\cos\theta)\cos(kr\sin\theta)\,r\,dr\,d\theta. \qquad ......(\mathrm{A}\,2)$$

The trigonometrical expressions can be written

$$\begin{aligned}&\cos(kr\cos\theta)\cos(kr\sin\theta)\\ &=\tfrac{1}{2}[\cos\{kr(\cos\theta+\sin\theta)\}+\cos\{kr(\cos\theta-\sin\theta)\}] \qquad ......(\mathrm{A}3)\\ &=\tfrac{1}{2}[\cos\{kr\sqrt{2}\cos(\theta-\tfrac{1}{4}\pi)\}+\cos\{kr\sqrt{2}\cos(\theta+\tfrac{1}{4}\pi)\}].\end{aligned}$$

The following three relations are also required:

$$\int_{\alpha=0}^{\pi}\cos(z\cos\alpha)\,d\alpha=\pi J_0(z); \qquad ......(\mathrm{A}\,4a)$$

$$\int_{\alpha=0}^{\pi}\cos(z\sin\alpha)\,d\alpha=\pi J_0(z) \qquad ......(\mathrm{A}\,4b)$$

$$\int zJ_0(z)\,dz=zJ_1(z) \qquad ......(\mathrm{A}\,5)$$

where $J_0(z)$ and $J_1(z)$ are Bessel functions of order zero and one. Using the equations (A3), (A4a) and (A5) the integral (A2) becomes $(\pi\sqrt{2}/k)\rho J_1(k\rho\sqrt{2})$. Similarly expression (A1b) in terms of polar coordinates is

$$2\int_{r=0}^{\varrho}\int_{\theta=0}^{\pi}[\cos(kr\sin\theta)+\cos(kr\cos\theta)]\,r\,dr\,d\theta.$$

Using the equations (A4a), (A4b) and (A5) this becomes $(4\pi/k)\rho J_1(k\rho)$.

REFERENCES

Bechmann, R., 1943, *Hochfrequenztech. u. Elektroakust.*, **61**, 1; 1951, *Proc. Phys. Soc.* B, **64**, 323.

Cady, W. G., 1946, *Piezoelectricity* (New York : McGraw-Hill).

Ekstein, H., 1944, *Phys. Rev.*, **66**, 108.

Kamayachi, Z., and Ishikawa, T., 1940, *Elektrotechn. J., Japan*, **4**, 243.

Mähly, H., 1945, *Helv. phys. Acta*, **18**, 248.

Mähly, H., and Trösch, A., 1947, *Helv. phys. Acta*, **20**, 253.

The Astigmatic Correction of Telescopes *

By R. S. LONGHURST †
Technical Optics Section, Imperial College, London

Read before the Optical Group 14*th December* 1950; *MS. received* 26*th October* 1950, *and in amended form* 30*th January* 1951

ABSTRACT. It is shown theoretically that if a telescope has positive field curvature the best possible image on the Gaussian plane is obtained by making the tangential field flat.

§1. INTRODUCTION

IN the usual terrestrial telescope the spherical aberration, coma, astigmatism, distortion and longitudinal and transverse chromatic aberrations may be corrected, but it is not possible to correct the Petzval field curvature. Thus, although it is possible to secure off-axis images of good quality, the image surface is not a plane. In systems where the field of view is small this error is of little consequence, but in many systems the field curvature assumes considerable importance. One of the most common telescope systems is that employed in the prismatic binocular, and this is usually called upon to present a wide angle of field.

In the case of the binocular telescope it is the practice of observers to focus on the axial image, the accommodation of the eyes being relaxed. That is to say, emmetropic observers adjust the instrument so that the final image is formed at infinity. Under these conditions the presence of positive field curvature causes the extra-axial image forming pencils to be convergent. It has been found (Wald and Griffin 1947) that observers are unable to make the muscular effort necessary to view such images. It will be appreciated that the effort which would be required in this case is directly opposite to that exerted in normal accommodation when the eye is made to bring divergent pencils to a focus in order to view real objects at finite distances.

If, after adjusting the focus of the instrument so that the axial image is at infinity, the observer views a marginal part of the field, the image perceived will be that formed on the Gaussian plane at infinity and the quality of the image will be poor in the presence of field curvature.

The conditions obtaining in such a system are most easily understood by considering the shapes of the emergent wave fronts. When the observer adjusts the focus of the instrument so as to view the axial image with relaxed accommodation the emergent wave front for an axial image will be plane, and the observer will have adjusted the focus of his eyes to receive plane waves. If the observer now views a marginal part of the field, the quality of the image perceived will depend on the deviation of the emergent wave front from the plane wave form. In the presence of positive field curvature alone, the disturbances from the marginal parts of the field will consist of convergent spherical waves and the deviation from the ideal wave front will be that of a sphere from its tangent plane.

---

* This investigation was included in work approved by the University of London for the degree of Ph.D.

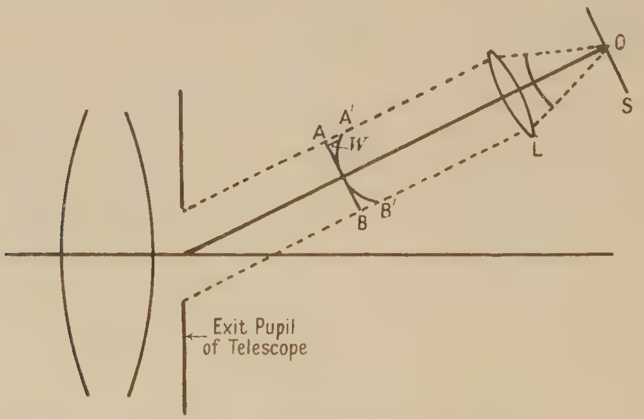
† Now at Chelsea Polytechnic, London.

One can exercise a limited control over the form of the emergent wave front by introducing negative astigmatism and consequently there may exist the possibility of improving the quality of the off-axis images. It will be appreciated that the introduction of coma will not assist in overcoming field curvature since the presence of coma does not materially move the best focus position longitudinally (Steward 1928). Moreover, the introduction of primary spherical aberration would affect the axial and extra-axial images equally and would do nothing to alter the relative positions of their best planes of focus.

As one introduces negative astigmatism, the wave front in both the sagittal and tangential sections approaches the plane wave form, but simultaneously the shapes of the wave front in the sagittal and tangential sections become more widely different, a condition which will impair the image quality. Alternatively one may say that initially the best focus will be situated on the Petzval surface which will be different from the gaussian plane. As one introduces negative astigmatism the sagittal and tangential image fields become more nearly flat but, since the distance between them increases, the best-focus image (which is located mid-way between the sagittal and tangential foci) deteriorates as it approaches the gaussian plane. The problem is, then, to determine the astigmatic correction which will lead to the most satisfactory image on the gaussian plane in the presence of positive Petzval field curvature.

§ 2. SOLUTION FOR SMALL ABERRATIONS

In the Figure AB represents a plane wave emerging from a telescopic system. The optical system of the observer's eye can be represented by the lens L and the retina by the screen S. If L is a perfect lens the plane wave AB will, after refraction, converge on to O where the usual Airy disc diffraction pattern will be formed on the screen. If the telescope system is not free from aberrations the emergent wave front will not be plane and the distribution of light on the screen will be modified. If distortion is present the point O is taken to be the point where the principal ray of the emergent pencil cuts the screen.



Let A′B′ represent a wave front with aberrations, and let *W* be the wave-front aberration, that is, the deviation from the plane AB. In the case of positive field curvature, the outer parts of the wave front will lie to the right of the reference surface AB.

Let E be the mean square deviation of the wave front from the reference surface. Then, by definition, $E=\iint W^2 d\omega$ where $d\omega = dA/A$ and A is the area of the exit pupil of the telescope. Any plane parallel to AB may be taken as the plane of reference and the value of E will vary according to the particular plane which is chosen. Let E_0 be the minimum value of E which is so obtained. Then it has been shown (Maréchal 1948) that, in the presence of small aberrations (i.e. for wave-front aberrations which do not exceed the Rayleigh quarter-wave limit when measured from the best reference surface), the intensity at O is given by e where $e \geqslant (1-2\pi^2 E_0/\lambda^2)^2$, e being taken as unity at the centre of a true Airy disc. That is, $e=1$ at O if a plane wave emerges from the telescope. Thus the intensity at O will be a maximum when the wave front has such a form that E_0 is a minimum. It may be noted that the condition for the best image irrespective of its location would be obtained by choosing as reference surface any sphere or any plane. The inability of the eye to exert negative accommodation restricts the position of focus to O or points to the right of O. This restricts the choice of reference surface to planes parallel to AB or spheres with their centres to the left of AB. The latter obviously lead to larger values of E and need not be considered.

Suppose the telescope system has both field curvature and astigmatism. Then the sagittal and tangential foci will be located on the principal ray in the region of O but neither necessarily at that point. Let s and t be the distances of these images from O. Maréchal has shown that, for primary aberrations,

$$E_0 = (u'^4/192)\{(t+s)^2 + 2(t-s)^2\}, \qquad \ldots\ldots(1)$$

where u' is the angular semi-aperture of the pencil converging on O; s and t are reckoned positive if measured from O to the right.

In the notation of Hopkins (1950) S_{III} and S_{IV} are the aberration coefficients associated with astigmatism and field curvature respectively. In terms of these coefficients

$$t = -\frac{1}{2u'^2}(3S_{\mathrm{III}} + S_{\mathrm{IV}}); \quad s = -\frac{1}{2u'^2}(S_{\mathrm{III}} + S_{\mathrm{IV}}), \qquad \ldots\ldots(2)$$

whence $E_0 = (6S_{\mathrm{III}}^2 + 4S_{\mathrm{III}}S_{\mathrm{IV}} + S_{\mathrm{IV}}^2)/192$.

In a telescope system, the value of S_{IV} is, in general, substantially constant, whereas the value of S_{III} can be varied over a wide range. The problem now is to decide upon the most suitable value of S_{III} for a given value of S_{IV}.

Writing $S_{\mathrm{III}} = -\gamma S_{\mathrm{IV}}$ one obtains the relation $E_0 = (S_{\mathrm{IV}}^2/192)(6\gamma^2 - 4\gamma + 1)$. The astigmatic correction giving the smallest value of E_0 and therefore greatest intensity at O is found by differentiating with respect to γ and equating to zero. One obtains in this way $\gamma = \frac{1}{3}$.

§ 3. LARGER ABERRATIONS

While the above analysis is sufficiently exact for small aberrations, it cannot readily be extended to include the larger aberrations that are frequently encountered in binocular telescopes. Coleman, Clark and Coleman (1947) have conducted a large number of experiments with visual instruments in which observers were asked to adjust the focus of the instruments for optimum image formation. The interferometer patterns were subsequently determined using a Twyman interferometer, and an analysis of the results indicated that observers adjust the focus of a telescope so that the interferogram is such that the largest possible circle can be inscribed in the space covered by a single light or dark interference area. If the diameter of this circle is expressed as a percentage of the diameter of the axial entrance pupil, the percentage is referred to as the Interferometer Quality, denoted by I.Q.

The region of the interferogram in which the I.Q. circle is inscribed corresponds to the area of the exit pupil over which the wave-front aberration does not exceed a quarter of a wavelength. The investigations were restricted to finding the position of the best focus in the presence of fixed aberrations, but they were not confined to systems having only small aberrations. Although these experiments did not take account of possible variations in the aberrations it is reasonable to suppose that the result holds generally and that the optimum state of correction is that which leads to the largest I.Q.

For any given state of correction the interferometer pattern can be calculated and the effect of varying the aberrations can be found. Thus it is possible to calculate the state of correction which gives the largest I.Q.

Consider a wave front exhibiting field curvature and astigmatism. Take as reference surface that plane which touches the wave front at the point of intersection of the principal ray. The wave-front aberration of a point on the wave front whose polar coordinates are (ρ, ϕ) is given (in Hopkins' notation) by

$$W=(\rho^2/R^2)[\tfrac{1}{2}S_{\mathrm{III}}\cos^2\phi+\tfrac{1}{4}(S_{\mathrm{III}}+S_{\mathrm{IV}})], \qquad \ldots\ldots(3)$$

where R is the value of ρ at full aperture and S_{III} and S_{IV} are the aberration coefficients for astigmatism and field curvature respectively, also at full aperture.

As before put $S_{\mathrm{III}}=-\gamma S_{\mathrm{IV}}$. Then $W=(\rho^2/R^2)S_{\mathrm{IV}}[-\tfrac{1}{2}\gamma\cos^2\phi+\tfrac{1}{4}(1-\gamma)]$ and for any reference plane parallel to the original plane but removed from it by a distance δ,

$$W=(\rho^2/R^2)S_{\mathrm{IV}}[-\tfrac{1}{2}\gamma\cos^2\phi+\tfrac{1}{4}(1-\gamma)]+\delta. \qquad \ldots\ldots(4)$$

We now have to find the value of γ which will give a region of the wave front throughout which $|W|\leqslant\lambda/4$ and in which the largest possible circle may be inscribed, it being assumed that the optimum value of δ is used. The value of $|\delta|$ cannot exceed a quarter of a wavelength since a larger value would imply the appearance of a new interference fringe at the centre of the pattern.

With rectangular axes x and y in the reference plane $x=\rho\sin\phi$, $y=\rho\cos\phi$, whence

$$W=(S_{\mathrm{IV}}/R^2)[\tfrac{1}{4}(1-\gamma)x^2+\tfrac{1}{4}(1-3\gamma)y^2]+\delta. \qquad \ldots\ldots(5)$$

If $\gamma<\tfrac{1}{3}$ the tangential image surface lies between the Petzval surface and the gaussian plane and the wave front lies wholly to the right of the reference surface. In this case the best reference plane is that which lies as far as possible to the right of the original plane, i.e. $\delta=-\lambda/4$. Then (from (5)) the locus of points on the wave front for which $W=+\lambda/4$ is

$$\frac{S_{\mathrm{IV}}(1-\gamma)}{2\lambda R^2}x^2+\frac{S_{\mathrm{IV}}(1-3\gamma)}{2\lambda R^2}y^2=1.$$

The diameter of the I.Q. circle is given by the minor axis of this ellipse, i.e. $[2\lambda R^2/(1-\gamma)S_{\mathrm{IV}}]^{1/2}$, and it is apparent that the largest I.Q. results when $1/(1-\gamma)$ has its largest value. Consequently, for the range $\gamma<\tfrac{1}{3}$ the I.Q. increases as γ tends to $\tfrac{1}{3}$.

If $\tfrac{1}{3}<\gamma<\tfrac{1}{2}$ it is necessary to rewrite equation (5) in the form

$$W=(S_{\mathrm{IV}}/R^2)[\tfrac{1}{4}(1-\gamma)x^2-\tfrac{1}{4}(3\gamma-1)y^2]+\delta.$$

The locus of $W=-\lambda/4$ has the equation

$$\frac{(3\gamma-1)S_{\mathrm{IV}}}{4R^2(\frac{1}{4}\lambda+\delta)}y^2-\frac{(1-\gamma)S_{\mathrm{IV}}}{4R^2(\frac{1}{4}\lambda+\delta)}x^2=1 \qquad \ldots\ldots(6)$$

and the locus of $W=+\lambda/4$ is

$$\frac{(1-\gamma)S_{\mathrm{IV}}}{4R^2(\frac{1}{4}\lambda-\delta)}x^2-\frac{(3\gamma-1)S_{\mathrm{IV}}}{4R^2(\frac{1}{4}\lambda-\delta)}y^2=1. \qquad \ldots\ldots(7)$$

Thus the first fringe of the interferogram takes the form of two hyperbolae.

The curve represented by (6) cuts the y axis at the point defined by

$$y^2=4R^2(\tfrac{1}{4}\lambda+\delta)/(3\gamma-1)S_{\mathrm{IV}} \qquad \ldots\ldots(8)$$

and the curve represented by (7) cuts the x axis at the point defined by

$$x^2=4R^2(\tfrac{1}{4}\lambda-\delta)/(1-\gamma)S_{\mathrm{IV}}. \qquad \ldots\ldots(9)$$

Now, as one varies δ the vertices of the hyperbolae move to and from the origin out of phase so that the largest I.Q. will result when the vertices are equidistant from the origin, i.e. when $4R^2(\frac{1}{4}\lambda+\delta)/(3\gamma-1)S_{\mathrm{IV}}=4R^2(\frac{1}{4}\lambda-\delta)/(1-\gamma)S_{\mathrm{IV}}$, from which δ is found to be given by $\delta=\lambda(\gamma-\frac{1}{2})/2\gamma$. This equation defines the optimum value of δ for values of γ in the range $\frac{1}{3}<\gamma<\frac{1}{2}$.

When δ takes this value the radius of the I.Q. circle is given (from (8) or (9)) by

$$r^2=R^2\lambda/S_{\mathrm{IV}}\gamma. \qquad \ldots\ldots(10)$$

Consequently the I.Q. increases as γ decreases and it follows that, in the presence of positive field curvature, the largest I.Q. results when $\gamma=\frac{1}{3}$.

This result has been confirmed experimentally for values of S_{IV} equal to 5, 10 and 15 wavelengths. The test objects employed were letters of the alphabet (high contrast) and a miniature portrait (low contrast).

§ 4. CONCLUSION

The results of § 2 and § 3 show that in order to obtain the best possible image on the gaussian plane in the presence of positive field curvature one should introduce negative astigmatism of amount given by the relation $S_{\mathrm{III}}=-\frac{1}{3}S_{\mathrm{IV}}$.

In other words, one should render the tangential image field flat. Although this result has been obtained with special reference to a telescopic system the result is perfectly general. The result is of particular importance in the case of binocular telescopes since this is a case when one cannot choose the plane of focus and one requires the best possible image on the gaussian plane.

ACKNOWLEDGMENTS

The author wishes to thank Professor L. C. Martin for his interest and encouragement and the British Scientific Instrument Research Association for financial assistance.

REFERENCES

Coleman, H. S., Clark, D. G., and Coleman, M. F., 1947, *J. Opt. Soc. Amer.*, **37**, 671.
Hopkins, H. H., 1950, *Wave Theory of Aberrations* (Oxford : Clarendon Press).
Maréchal, A., 1948, *Thesis*, Paris.
Steward, G. C., 1928, *The Symmetrical Optical System* (Cambridge : University Press).
Wald, G., andGriffin, D. R., 1947, *J. Opt. Soc. Amer.*, **37**, 321.

LETTERS TO THE EDITOR

Plasma Oscillations in Meteor Trails

Experiments have recently been carried out to investigate the scattering of radio waves from the columns of ionization produced by meteors during their passage through the upper atmosphere.

This problem has been discussed theoretically by Lovell (1947) and by Lovell and Clegg (1948), who derived an expression for the power scattered from a meteor trail of small diameter by treating the electrons as individual free scatterers. It has been suggested by Herlofson (1948, 1951), however, that if the electron density in the column is of the order of the critical density for the wave frequency used, this theory will only apply when the electric vector is parallel to the axis of the trail; and that in cases when the incident radiation is polarized in a direction perpendicular to the axis, the amplitude of the echo may be considerably enhanced by plasma resonance effects. Such effects have been observed experimentally in discharge tubes by Denno, Prime and Craggs (1950), and by Romel (1951).

The present experiments employed the standard radio echo technique (see Lovell, Banwell and Clegg 1947). The essential features of the apparatus were twin directional aerials, each operating on a wavelength of 4 metres and capable of producing two identical beams of radiation, with planes of polarization mutually perpendicular. Each aerial was common to the transmitter and the receiver, and radio-frequency pulses from the transmitter were radiated simultaneously from both with a recurrence frequency of 600 c/s., but the two were switched during reception, so that echoes were received alternately on each polarization. The apparatus was designed for use during periods of intense meteor shower activity from specific radiants. The aerials could then be adjusted so that their electric vectors were respectively longitudinal and transverse relative to the axes of the meteor trails. The echoes were photographed by an automatic recorder of the type described by Davies and Ellyett (1949).

Measurements were made during the Geminid meteor shower of December 1948, and photographic records of 426 echoes were obtained. It was found that the echoes of duration less than 0·15 second showed somewhat different characteristics from those of longer duration, and it is convenient to consider the two groups separately.

Short duration echoes. A total of 378 echoes of duration less than 0·15 second was observed; 371 appeared at greater amplitude on the transversely polarized aerial (i.e. when the electric vector was perpendicular to the axis of the trail), and 201 of these were not observed at all on the longitudinally polarized aerial. Of the remaining 170 which showed measurable ratios, the amplitudes observed on the transversely polarized system exceeded those of the longitudinal by factors ranging from 1 to 12. These ratios seldom changed by more than 25% during the lifetime of the individual responses. The frequency distribution of the observed mean ratios is shown in the histogram in Figure 1.

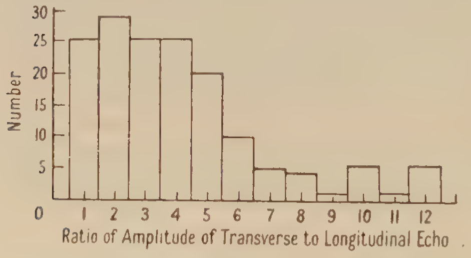


Figure 1.

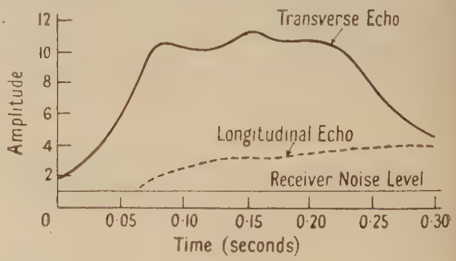


Figure 2.

These results are subject to discrimination in favour of ratios close to unity, since the probability of detecting the longitudinal component of an echo decreases with increasing ratio. A statistical analysis indicates that the true number of short duration echoes showing ratios greater than 10 to 1 probably exceeds the measured number by a factor of the order of 3, and that in approximately 15 cases the ratio was greater than 12 : 1.

Long duration echoes. Of the 48 echoes of duration greater than 0·15 second, all were observed on both polarizations. An interesting feature of echoes of this group was the variation in the ratio with time. Thus in the majority of cases the amplitude of the transverse echo initially exceeded that of the longitudinal by a factor of 3 or more, but the latter generally tended to increase during the lifetime of the echo, as indicated in Figure 2.

These observations appear to confirm the occurrence of plasma resonance in the majority of the meteor trails detected during the period of observation with this apparatus. The relationship of these results to contemporary scattering theory will be discussed later.

The University of Manchester,
Jodrell Bank Experimental Station,
Holmes Chapel, Cheshire.
21st May 1951.

J. A. Clegg.
R. L. Closs.

Davies, J. G., and Ellyett, C. D., 1949, *Phil. Mag.*, **40**, 614.
Denno, S. N., Prime, H. A., and Craggs, J. D., 1950, *Proc. Phys. Soc.* B, **63**, 726.
Herlofson, N., 1948, *Observatory*, **68**, 230; 1951, *Ark. Fys.*, **3**, No. 15, 247.
Lovell, A. C. B., 1947, *Nature, Lond.*, **160**, 670.
Lovell, A. C. B., Banwell, C. J., and Clegg, J. A., 1947, *Mon. Not. Roy. Ast. Soc.*, **107**, 164.
Lovell, A. C. B., and Clegg, J. A., 1948, *Proc. Phys. Soc.*, **60**, 491.
Romel, D., 1951, *Nature, Lond.*, **167**, 243.

Multiple Traversal of Cyclotron Targets: an Extension to the Theory

The particles accelerated in high energy cyclotrons may pass through an internal target more than once. In an earlier communication (Cassels, Dickson and Howlett 1951) the effective thickness T of the target, measured in radiation length units, was calculated as a function of $\lambda = w^2nd^2/4tr_0^2$. Here t is the actual thickness of the target, also in radiation length units, and w, n, d and r_0 are parameters determined by the operating conditions of the cyclotron.

The discussion was restricted to values of $\lambda \leqslant 4$, so that the theory could not be applied to very thin targets. Recently, however, both Dr. B. Davison and Dr. T. H. R. Skyrme of this Establishment have independently applied Fourier transform techniques to the integral equation (5) of the earlier paper. Although they have used rather different methods, both have arrived at the same form for the asymptotic variation of T when λ is large. The analysis is fairly complicated, but Dr. Davison's solution can be written in the simple form,

$$4r_0^2T/w^2nd^2 = 2 + 1{\cdot}6480\lambda^{-1/2} + 0{\cdot}3395\lambda^{-1} + O[\lambda^{-1}\exp\{-2(\pi\lambda)^{1/2}\}].$$

The following results have been derived by taking the first three terms of this expression.

λ	1	2	3	4	5	10	20	50	100
$4r_0^2T/w^2nd^2$	3·99	3·34	3·06	2·91	2·80	2·56	2·39	2·24	2·17

Comparison of these figures with those given earlier shows that for $\lambda \geqslant 4$ accurate results have been obtained by this procedure.

Since the effective thickness of the target becomes constant for very large λ, it is clear that very high specific activities can be obtained by bombarding thin targets. This has also been pointed out by Knox (1951).

We are very much indebted to Dr. Davison and Dr. Skyrme for their work on this problem.

Atomic Energy Establishment,
Harwell, Didcot, Berks.
30th May 1951.

J. M. Cassels.
J. M. Dickson.
J. Howlett.

Cassels, J. M., Dickson, J. M., and Howlett, J., 1951, *Proc. Phys. Soc.* B, **64**, 590.
Knox, W. J., 1951, *Phys. Rev.*, **81**, 693.

REVIEWS OF BOOKS

Physics, by S. G. Starling and A. J. Woodall. Pp. xvi + 1301. 1st Edition. (London : Longmans, Green, 1950.) 42*s*.

In an age of small monographs and of textbooks of composite authorship by teams of specialists this comprehensive treatise, dealing with physics in all its major branches, makes a welcome appearance. Those of us who long ago were brought up on such textbooks as Ganot, Watson, and Poynting and Thomson know full well the advantage of presentation of all parts of the subject by the same teachers whose outlook, method and style are apparent throughout : the combination we have here of an author renowned as a teacher, writer and exponent of his subject and a modern virile senior lecturer has proved particularly happy. The simple title *Physics* is the right one.

Wisely, this book is not set rigidly to any prescribed syllabus but presents basic physics, at advanced level, as seen by the authors. It will be of real value to the student reading for a general B.Sc. degree or taking a two-year ancillary course or working for special physics, particularly Part I under the new regulations of the University of London : such students will experience little trouble with the mathematics involved.

Some 250 pages are given to General Physics and Sound, while Heat and Light in 680 pages receive most generous treatment. Page 927 is reached before Electricity begins, and thus this large portion of the subject, including Atomic Physics, seems relatively cramped in some 300 pages. The early chapters on harmonic motion and gravitation are very thorough while those on elasticity, surface tension and viscosity are clear and adequate. The one conspicuous omission in this section is any treatment of moments of inertia, a subject rarely completed in an intermediate science course. The wide separation of the Sound from the General Physics is to be regretted, delaying, as it does, the chapter on wave motion (which appears long after physical optics has been treated) and necessitating consideration of forced vibration in both chapters 2 and 36.

The section on Heat is well arranged and is made unusually interesting, the chapters on kinetic theory, vapours, equations of state, solutions and the second law of thermodynamics being particularly impressive. This section closes with a chapter on statistical mechanics and quantum theory which, perhaps inevitably, seems at a level well beyond that set for the book as a whole. Jaeger and Diesselhorst's classical measurements of thermal conductivity might have had fuller treatment, showing temperature distribution expressed in terms of potential rather than distance, and one might have hoped for some detail on modern methods of measuring the mechanical equivalent of heat.

The Optics section is outstanding. The comprehensive treatment of refraction at a single surface and by a prism is good : mirrors and lenses are treated not only conventionally but by power and sagitta methods. Sign conventions are considered and frequently compared, while the useful discussions on aberrations are illustrated quantitatively by numerical data. The excellent chapters on diffraction, polarization and spectroscopy are beautifully illustrated by helpful plates and diagrams. With more care, however, given to logical arrangements space might have been saved : concepts, e.g. aplanatic surfaces, are frequently introduced when wanted and discussed much later. Angular magnification is not mentioned until the telescope is reached.

The electrical section of the book, with the exception of excellent chapters on electrolysis and magnetization of materials, seems too condensed. One might have hoped to see a short account of the Schuster magnetometer, methods of demonstrating the Peltier and Thomson effects, a vector diagram of the air transformer and an outline of the Campbell or Astbury standard of mutal inductance on which so many electrical measurements depend. The diagram (Figure 48, 8) of the N.P.L. current balance needs modification for it is not, as stated, of the Rayleigh type but of the Ayrton–Jones pattern. Perhaps the least happy portion of the book is that dealing with alternating current bridges : from statements, repeatedly made, it is clear that the authors are under the impression that if the conditions for balance of an A.C. bridge are independent of frequency the bridge may be used with ballistic galvanometer and battery. This of course is not the true criterion of such dual use. The Owen bridge, unlike the Anderson, is completely insensitive to inductance at zero frequency and, like the Schering, when balanced ballistically requires only the fulfilment of the De Sauty relation. As the Wien bridge is given, its use for measuring power factor or frequency might have been emphasized, but a brief description of the condenser series bridge or the Schering bridge for measuring phase defect of condensers would have been more useful.

The section on Atomic Physics contains very much that is important and useful to students but appears to lack logical sequence. Under gases, electrons, etc., an article on deuterium is followed by others on the cathode-ray oscillograph, the electron microscope, time bases and radar, but electronic valves are dealt with much later under atoms and radiation in between the photo-effect and a discussion on relativity, while the magnetron appears much earlier. X-rays are treated too briefly and no modern methods of crystal analysis are given. The discovery of the neutron and its properties appear on p. 1215 but artificial radioactivity on p. 1187. Only Millikan's discarded theory of cosmic rays is given, and one is led to believe that little is known about meson masses.

It is of course very easy for a specialist on any branch of the subject to suggest additions or deletions in a book of this nature or to criticize the sequence of presentation. Taken as a whole the book is really good and useful. The explanations are lucid, the type is clear and the diagrams and photographic plates, many of which have been specially prepared, are excellent. Tables of logarithms, trigonometrical functions, physical constants and principal symbols are given as well as a useful collection of examination questions which might with advantage have been dated. Reference to any part of the subject is facilitated by a well-arranged, chaptered and paged table of contents together with a good index. This portmanteau work, the best of its kind in existence, was clearly " not ordained to be gazed upon or to be carried about " but is designed for use in the home and the library. The labour entailed in its preparation has been very great and has, with a world war intervening, been spread over many years. The authors are to be congratulated on their courage in undertaking and completing this large textbook, which will be of inestimable benefit to the student and teacher.

H. R. NETTLETON.

Physical Mechanics: an Intermediate Text for Students of the Physical Sciences, by R. B. LINDSAY. Pp. x + 451. 2nd Edition. (New York: Van Nostrand; London: Macmillan, 1950.) $5; 37*s.* 6*d.*

It would be an interesting experiment first to compile a list of textbooks on mechanics, published, say, during the last fifty or sixty years, and then to take a vote amongst the Fellows of the Physical Society in order to discover which is the most popular. If members were able to shed their prejudices in favour of the books whence they derived their undergraduate nurture, and if the difficulty of cost, due to the present-day value of the pound sterling, could be surmounted, the book under review would undoubtedly be high on the list, partly because of its high quality and partly because it is a book for physicists written by a physicist. There is little difficulty in drawing up a long list of books on mechanics written from the standpoint of mathematicians, and there is no lack of textbooks designed to illumine the minds of engineers. Physicists however do not seem to take kindly to the task of catering for their kind in this matter, but Professor Lindsay is an exception, and he has succeeded in fulfilling his aim of stressing " the fundamental concepts and principles of mechanics and their use in all branches of physics ". The book treats the statics and dynamics of particles, systems of particles, rigid bodies, deformable solids and fluids, and it includes chapters on constrained motion, oscillation and wave motion. The concluding chapter treats Hamilton's principle, Lagrange's equations and also contains a short section on wave mechanics.

The treatment is clear and lucid throughout, and the examples used to illustrate theorems are chosen for their physical rather than mathematical interest. One excellent feature of the book is the collections of problems, and a student who can successfully tackle these will have assimilated a thorough knowledge of the basic principles of mechanics.

R. M. D.

The Structure and Mechanical Properties of Metals, by BRUCE CHALMERS. (Monographs on Metallic Materials, Vol. 2.) Pp. ix+132. (London: Chapman and Hall, 1951.) 18*s*.

The users of metals and alloys are intimately concerned with the mechanical properties of the materials which form the basis of their manufactures. In particular, designers of machines made from metal must know, and be able to compare and contrast, the properties of the materials available to them. There exists, therefore, a very great deal of knowledge, much of it empirical, about how materials behave under a variety of conditions, but it would not be unfair to suggest that knowledge about why metals and alloys behave as they do is much less widely disseminated. Those who have to attempt to work out the effects of changes in the composition or the heat treatment of their metallic materials, such as may be imposed by economic considerations, may find it a difficult task without this knowledge. To keep abreast of modern tendencies—jet engines, supersonic flight, and the like—the designer may call desperately for alloys having certain specific properties, often without any clear idea of how these properties may be developed, or indeed of whether they can be developed at all. Of course, these questions could ultimately be settled by experimental investigation, but the work is likely to be a long and costly undertaking unless it is based upon a sound understanding of the general factors affecting the mechanical properties of metals and alloys. Although no one would claim that there is any complete theory, much progress has been made, particularly by the physicists. Professor Chalmers' book may be described as setting out, in a most agreeable and readable way, some of the most important factors which control mechanical properties. He has aimed at providing the non-specialist reader with the simplest possible picture of the structure of metals and alloys and its relation to their mechanical properties, and he has hit his mark. Mathematics has been miraculously avoided, and this has necessitated a rather superficial approach to some aspects of the subject. The consequent over-simplification and the occasional half-truth are deliberate, and the author makes no apology for them. The reader will realize, as he progresses with the book, that no apology is needed; rather should he congratulate Professor Chalmers on having set the correct standard for the readers to whom his book is addressed, and on having adhered to this standard throughout.

The book begins with a description of the structures of pure metals, the cohesive forces in metals, and the manner in which metals crystallize from the molten state. The structures of alloys are considered next, and the subject of constitutional diagrams is adequately introduced, with typical examples. In Chapter 3 the deformation of metals is discussed, proper attention being paid to the fundamental physics of the processes involved. This is followed by an account of the manner in which heat treatment modifies the structures of alloys, and here phase changes, diffusion effects, diffusionless transformation, recovery and recrystallization all receive consideration. After a short chapter on methods available for examining the structures of metals and alloys, there is an excellent account of the manner in which the structural characteristics previously explained affect the mechanical properties of technical importance. Due emphasis is placed on the interpretation of mechanical tests, and several points with regard to which some confusion has existed in the past are well explained. This chapter summarizes excellently the work which has been done on relating structure and properties.

A pleasing feature is that the book is really up-to-date. An account of recent work on the raising of Young's modulus of aluminium alloys, which may prove important to aircraft designers, is included, while throughout the book the concept of the dislocation is used in discussing plasticity. In this book, difficult subjects lose much of their terror for the non-specialist, and the treatment will be especially appreciated by this class of reader. The book is assured of a hearty, and a widespread, welcome; both the author and the Council of the Royal Aeronautical Society are to be congratulated.

There are some minor blemishes which should have been eliminated by the proof-reader. Misprints are a little too frequent—magnesium melts at 649° C. and not 659° C. as stated, while occasionally the descriptive lettering on diagrams does not correspond with the textual discussion. Illustrations are well chosen and in general excellent, but are at times confusing. Thus, one diagram is intended to illustrate in four stages how a moving dislocation produces slip; the stages are not labelled to correspond with the text, and the whole diagram appears to have been inverted. Occasionally the text refers to the wrong illustration. Such imperfections are the more serious the less the reader already knows of the subject, and it is hoped that they will be rectified in the later editions which are sure to be demanded.

This book may be recommended to all who seek a clearly written, authoritative introduction to the subject. G. V. RAYNOR.

Semiconductors, by D. A. WRIGHT. (Methuen's Monographs on Physical Subjects.) Pp. viii + 130. (London: Methuen, 1950.) 7*s.* 6*d.*

After two introductory chapters on electrons in metals and in crystals (50 pages) the author deals in the six following chapters with applications. Two of them (20 pages) are connected with problems of conductivity, the rest (78 pages) are devoted to emission of electrons from free surfaces. These latter chapters are well written and lead the reader up to recent developments; they contain a fair number of references to the original literature. Other aspects of semiconductors are, in comparison, treated in very little detail. This holds in particular of the basic questions such as structure, mean free path and others. No derivation is given for the formula describing the temperature dependence of the number of conduction electrons, whereas the question of the penetration of electrons through potential barriers is treated in some detail. It appears that the author intended the book to be mainly an introduction to the emission of electrons from surfaces. He has well succeeded in this respect. It would have been of benefit to the prospective reader if this restriction in the scope of the book had been indicated by an appropriate subtitle. H. FRÖHLICH.

Fundamentals of Optics, by F. A. JENKINS and H. E. WHITE. Pp. xi + 647. 2nd Edition. (London and New York: McGraw-Hill, 1950.) 50*s.*

The first edition of this book was well known in this country under the title *Fundamentals of Physical Optics* and has proved a valuable textbook for physics students. The present volume, however, is much more than a simple revision of its predecessor, for the change in title is a reflection of a significant change of the book's scope and contents. The earlier text has, indeed, been revised and brought up to date, but ten chapters on geometrical optics have been added, and also a final chapter on the quantum nature of light.

The book is very clearly written and is produced at the very high standard which we have come to associate with American publications. The great activity in optics which has been a characteristic of recent years has enabled the authors to use modern applications to illustrate well-established principles, so that the reader becomes conscious that he is studying a live and exciting subject. This is not always a feature of books on optics.

To the reviewer, the value of the new geometrical optics section is to some extent diminished by the use of a non-Cartesian sign convention, in which positive distances are measured to the left of the lens in the object space and to the right of the lens in the image space, on the assumption that the light is travelling from left to right. The merits of different sign conventions are admittedly affected by one's familiarity with them, but this is not the sign convention the reviewer would choose.

On page 108 a function described as the 'normal magnification' of a telescope or microscope seems to be much too closely related to the diameter of the eye pupil to be of much significance, since the pupil size is a very variable quantity, and in many cases the exit-pupil of the instrument is very much smaller than the eye pupil. Another minor item for criticism is the use of simple lenses on pages 166 and 167 to illustrate the astronomical telescope; these are certainly out of keeping with the illustrations of microscope objectives and camera lenses in the same chapter. Again, the compensator described on page 252 is not the type usually associated with the Rayleigh refractometer.

In a book such as this, where there is no hesitation to discuss some of the most difficult conceptions about light, there is sometimes a surprising tendency to omit the analysis of less difficult, but quite important, matters. An example of this is the discussion of the intensity distribution of the fringes in the Fabry and Perot interferometer : the distribution is illustrated graphically, but the formula itself is neither derived nor quoted. Other sections in which the reader's palate is tickled, but hardly satisfied, are those dealing with the phase-contrast microscope and with fluorescence.

Yet the overall impression is that this is a sound and stimulating book, and to ask for more in a volume which already exceeds 600 pages is perhaps a tribute rather than a criticism.

W. D. WRIGHT.

Ultrasonics, by P. VIGOUREUX. Pp. 163. (London : Chapman and Hall, 1950.) 25*s*.

When a sound wave travels through a medium the energy of any given volume is constantly changing from kinetic to potential, and vice versa, at the wave frequency. If, for any reason, there is a time delay in the interchange of a part of this energy then the wave will be absorbed, and both the absorption and velocity will vary characteristically with the frequency. In many gases there is such a delay associated with the vibrational contribution to the specific heat. This relaxation process is now believed to occur in certain liquids such as benzene and carbon disulphide, whose high absorption of ultrasonic waves has long been a problem. It seems very likely that there are, in other liquids, further relaxation phenomena connected with the structure of the liquid which can also promote absorption and dispersion. In the last five or six years considerable progress has been made in the study of ultrasonic propagation in liquids, progress comparable to that made in the decade before the war in investigating gases. This progress has been due very largely to wartime work on radar, whose pulse technique has offered a new and powerful method of measuring the velocity and absorption of ultrasonic waves in liquids. For gases, however, the method is of little value because of the difficulty of securing a wide band coupling into a gas from the transducer. If this new method of studying the behaviour of liquids is vigorously pursued it may well provide a powerful tool for solving some of the outstanding problems of the liquid state.

In this book Dr. Vigoureux has most lucidly summarized everything now known about the physical aspect of propagation in liquids and gases and the experimental techniques appropriate to its investigation. The book is divided into five main chapters headed Generation, Propagation, Observation, Gases, Liquids, in which, by omitting all technical matter, the author has given a complete and up-to-date survey of the subject in a remarkably small compass. In order to give the clearest picture of the more fundamental features of propagation the author has deliberately either omitted or only outlined such topics as the optical methods of measurement and demonstration, the effects of waves of large amplitude, and transmission through boundaries and in solids. The essential results in the theory of the relaxation process are derived without recourse to complex mathematics, their physical significance is explained, and the whole discussion of propagation in liquids and gases is carried out in terms of that process. Electrical circuit analogies are introduced to help to explain the relaxation process and the matching of the transducer to the medium and to the transmitter. The treatment is such as should appeal equally to the physicist and to the electrical engineer.

There is a large number of diagrams in the text, and tables of values for the velocity and absorption are given for the principal gases and liquids and for certain mixtures. The printing and reproduction of diagrams are a model of clarity. A short mathematical appendix on the impedance of the transducer in the interferometer is followed by a full bibliography covering the period 1939 to 1950 and name and subject indexes to the text. Dr. Vigoureux's authoritative monograph is sure of a good reception from physicists and engineers in this country as it is the first book on ultrasonics to be written by an English author, and is in fact only the second in English on this subject published since 1938. Although it cannot be compared for completeness of detail with the 5th edition of Professor Bergmann's *Der Ultraschall* published in 1949, this book is the best introduction to ultrasonics now available. It is to be hoped that, if the author prepares a revised edition, the subject will have advanced sufficiently to permit the inclusion of work on solids as well as liquids and gases.

J. M. M. P.

An Introduction to Servo-Mechanisms, by A. Porter. (Methuen's Monographs on Physical Subjects.) Pp. vii + 154. (London: Methuen, 1950.) 7*s.* 6*d.*

With this small volume, written by an author who has expert knowledge of his subject and has made substantial contributions to its advancement, another useful member has been added to the series of Methuen Monographs. Within the last few years servo-mechanisms have come to the fore in a wide variety of applications. This has been to a large extent due to the rapid development of electronic techniques and their applications to radar and other military equipment. The widespread application of these techniques to industrial processes is a more recent development and is becoming of extreme importance as the need for accuracy of control becomes increasingly greater.

An extensive literature has already grown up around the subject and the time is ripe for a brief review of the subject, including a guide to the literature. It is most important that a beginner or non-specialist reader should get at the start a full picture of the most important aspects of the subject, together with an introduction to the more important papers which constitute the milestones of its development. This is admirably achieved in the present monograph. It introduces in logical sequence the fundamental principles, with examples showing their practical application.

Beginning with the general principles of feedback monitoring, the author develops the basic equations of linear servo systems. He than treats in turn transient and harmonic response of servos, introducing the various graphical and other methods used to study their characteristics. The important problems associated with the stability of servo systems are introduced and the volume ends with a short account of non-linear systems.

The book is not intended to provide an engineer with data for design of a servo system to meet the needs of any particular problem, but rather as an introduction of the subject to the non-specialist reader. The author uses throughout a mathematical approach, and, although no very difficult mathematics is used, the reader must be prepared to follow a mathematical type of argument throughout. This may be found to be a drawback by some readers. The volume is remarkably free from printing or other errors and can be fully recommended for the purpose for which it is intended. R. A. SMITH.

Operational Calculus based on the Two-sided Laplace Integral, by B. van der Pol and H. Bremmer. Pp. xiii + 415. (London : Cambridge University Press, 1950.) 55*s.*

This is by no means an elementary book, nor is it intended to show directly how the operational calculus is to be applied to given problems in electrical circuit theory or in diffusion problems. There are several good textbooks available on these matters, but this one is devoted to the pure mathematics of the subject. The word ' two-sided ' in the title refers to the range of integration in the basic integral (Mellin or Fourier), a one-sided integral ranging from zero to infinity, but a two-sided one from negative to positive infinity.

Apart from discussing the basis of the calculus, the authors consider its application to partial differential equations, difference equations and integral equations, but in many ways the most interesting matter which crops up is the realization that it can throw light on the theory of numbers. Riemann's zeta function presents itself quite naturally in the development of the theory, and a number of theorems on prime numbers and divisibility can be deduced rather more easily by the methods of the operational calculus than by other methods.

One is glad to see that the authors have a great respect for the pioneer work of Heaviside, despite its experimental and non-rigorous character. J. H. A.

The Nomogram, by H. J. Allcock and J. R. Jones, revised by J. G. L. Michel. Pp. x + 238. 4th Edition. (London : Pitman, 1950.) 18*s.*

To the uninitiated there is always an element of magic in a nomogram, with its strangely graduated, inclined and curving axes, which performs computations so effortlessly. It reminds one of the boy calculating prodigy who, when asked how he found his answers replied : " God puts them into my head, but I cannot put them into yours." The authors of this book show us all how to become calculating prodigies. Their exposition of both theory and practice is so systematic and lucid that success is assured to all who have the will to try.

In his original presentation D'Ocagne presented the subject in terms of projective geometry and parallel line coordinates, with the result that many people were trying to learn something new in terms of the unfamiliar. The present exposition avoids this pitfall, being based for the most part on cartesian coordinates and determinants. The reviser has introduced parallel line coordinates in the last chapter, where he unifies alignment and intersection nomograms.

We are introduced to the classification of nomograms by classes and genera, and by learning to associate a given formula with its typical nomogram we soon march confidently in unfamilar territory, forming, first, the basic determinant, then the constructional determinant and, finally, the nomogram itself.

There is an interesting chapter on miscellaneous forms of nomograms, including polar, hexagonal and circular types, and a useful and instructive appendix gives a summary of the standard forms of the nomogram of the third class (i.e. for formulae involving three variables). Both author and publishers are to be congratulated on this revised version of a well-known book, which should be carefully studied by all interested in this work, whether from a theoretical or a practical standpoint. M. J. M.

CONTENTS FOR SECTION A

PAGE

Dr. M. Blackman. On θ Values in the Resistance of Metals 681

Mr. M. Smollett and Dr. M. Blackman. Lattice Stability and Oriented Overgrowth 683

Prof. L. F. Bates and Mr. G. W. Wilson. Some New Bitter Patterns on a Single Crystal of Nickel 691

Mr. Suraj N. Gupta. On Stueckelberg's Treatment of the Vector Meson Field . 695

Mr. R. J. Blin-Stoyle. Polarized Nuclear Reactions 700

Dr. S. A. Kushneriuk and Dr. M. A. Preston. The Scattering Phase Shifts of Two-Nucleon Systems 712

Mr. H. Messel. Average Numbers for the Development of the Nucleon Component of the Cosmic Radiation in Air 726

Dr. E. P. George and Dr. P. T. Trent. Penetrating Bursts of Cosmic Radiation . 733

Mr. J. W. Sturgess. Observations of the Latitude Effect of Penetrating Showers . 745

Letters to the Editor :

- Dr. R. P. Hudson. Investigation of Superconductivity in Lead Compounds . 751
- Dr. E. Billig. Effect of Temperature on the Height of Potential Barriers and on the Breakdown Voltage of Contact Rectifiers 752
- Mr. W. B. Mims and Dr. H. Halban. K Capture in ^{128}I 753
- Mr. A. Ashmore and Mr. J. F. Raffle. Magnetic Analysis of α-Particle Groups from (d, α) Reactions 754
- Mr. G. A. Jones and Dr. D. H. Wilkinson. The Reaction ^{13}C(αn)^{16}O. . 756
- Mr. F. A. El-Bedewi, Mr. R. Middleton and Mr. C. T. Tai. Angular Distributions of the Two Long-Range Neutron Groups of the Reaction ^{16}O(d, n)^{17}F 756
- Dr. B. Bleaney, Mr. K. D. Bowers and Dr. D. J. E. Ingram. Quadrupole Moments of Copper 63 and 65 758

Reviews of Books 760

Contents for Section B 761

Abstracts for Section B 762

ABSTRACTS FOR SECTION A

On θ Values in the Resistance of Metals, by M. BLACKMAN.

ABSTRACT. The θ value to be expected on the Bloch theory of the temperature dependence of the resistance of metals is shown to be much larger than that observed experimentally.

Lattice Stability and Oriented Overgrowth, by M. SMOLLETT and M. BLACKMAN.

ABSTRACT. The elastic constants of a two-dimensional ionic lattice, strained to fit on to the cube face of a substrate of the rocksalt type, are worked out. From these constants the stability conditions of the strained layer are deduced. These conditions are studied numerically in a number of selected cases where oriented overgrowth is observed in experiment. The calculations lead to conclusions which are in disagreement with a suggestion of Frank and van der Merwe as to the general mechanism of oriented overgrowth.

Some New Bitter Patterns on a Single Crystal of Nickel, by L. F. BATES and G. W. WILSON.

ABSTRACT. Some new Bitter figures have been obtained on a single crystal of nickel whose surface is a (111) plane when magnetized along a [110] direction, and a curve obtained of domain boundary spacing as a function of the effective magnetic field. Patterns of a complex nature, probably due to closure domains, were also obtained on a (112) plane.

On Stueckelberg's Treatment of the Vector Meson Field, by SURAJ N. GUPTA.

ABSTRACT. Stueckelberg's treatment of the vector meson field is investigated, and is found to involve difficulties similar to those of Fermi's treatment of the radiation field. It is shown that these difficulties may be removed by the same procedure as has already been applied to the radiation field in an earlier paper, which involves the use of an indefinite metric and a modification of the supplementary condition.

Polarized Nuclear Reactions, by R. J. BLIN-STOYLE.

ABSTRACT. Using properties of states of quantized angular momentum, general formulae are obtained for the angular distribution in nuclear reactions and scattering processes in which the bombarding particles and target nuclei are polarized. A simple symmetrical formula for use in specific calculations is given. The following cases are considered in detail : (*a*) polarized bombarding particles only, (*b*) target nuclei polarized in a magnetic field and bombarded by polarized particles. Finally, general predictions are made concerning the polarization of particles resulting from any reaction or scattering process and the results applied to the D–D reaction which is suggested as a source of fast polarized neutrons and protons.

The Scattering Phase Shifts of Two-Nucleon Systems, by S. A. KUSHNERIUK and M. A. PRESTON.

ABSTRACT. Energy expansions for simple functions of the scattering phase shifts are given for all relative angular momenta for central forces in neutron–proton and proton–proton scattering and for non-central static forces in proton–proton scattering including the coupling of P and F states. The differential scattering cross section for this latter case is given in an appendix.

Average Numbers for the Development of the Nucleon Component of the Cosmic Radiation in Air, by H. Messel.

ABSTRACT. Comprehensive results are presented in graphic form for nucleons with energies equal to or greater than 500 Mev., based on calculations using a new cross section for nucleon–nucleon collisions and in the light of recent experimental evidence. The results obtained are in good agreement with existing experimental observations and it is felt that the high-energy nucleon component of the cosmic radiation can now be satisfactorily accounted for by our theory. The basic assumptions underlying the theory are discussed.

Penetrating Bursts of Cosmic Radiation, by E. P. George and P. T. Trent.

ABSTRACT. Bursts in an ion chamber, accompanied by the discharge of one or more of 18 surrounding Geiger counters separated from it by 15 cm. Pb, have been studied at the Jungfraujoch, at sea level and below ground at a depth of 60 m. water. With 15 cm. Pb over the chamber the bursts at altitude were separated into those associated with air showers (accompanied bursts) and the remainder (local bursts). The accompanied penetrating bursts appear to be associated with the incidence of high density regions of air showers. The local penetrating burst size increases with the number of penetrating particles recorded consistent with their interpretation in terms of mixed showers involving the simultaneous production of charged and neutral mesons in energetic nuclear collisions of the N-component.

Underground, about 10% of the bursts under 15 cm. Pb are of the mixed shower type. These underground bursts are attributed to nuclear excitations produced by fast μ-mesons, resulting again in charged and neutral meson emission. About one-third of the bursts observed below ground with no Pb over the chamber appear to be produced by single penetrating particles which release, by an unidentified process, 20 Mev. and more in the ion chamber.

Observations of the Latitude Effect of Penetrating Showers, by J. W. Sturgess.

ABSTRACT. The latitude effect of the local penetrating showers of cosmic radiation produced in paraffin has been measured at sea level and found to be in the ratio of $1{\cdot}8 \pm 0{\cdot}4$ between London and Dakar, French West Africa. The extensive penetrating showers showed no latitude effect, which is consistent with the greater energy of these showers. The background frequency of local showers, i.e. those recorded with no absorber above the counter system, was found to be insensitive to latitude and reasons for this are discussed.

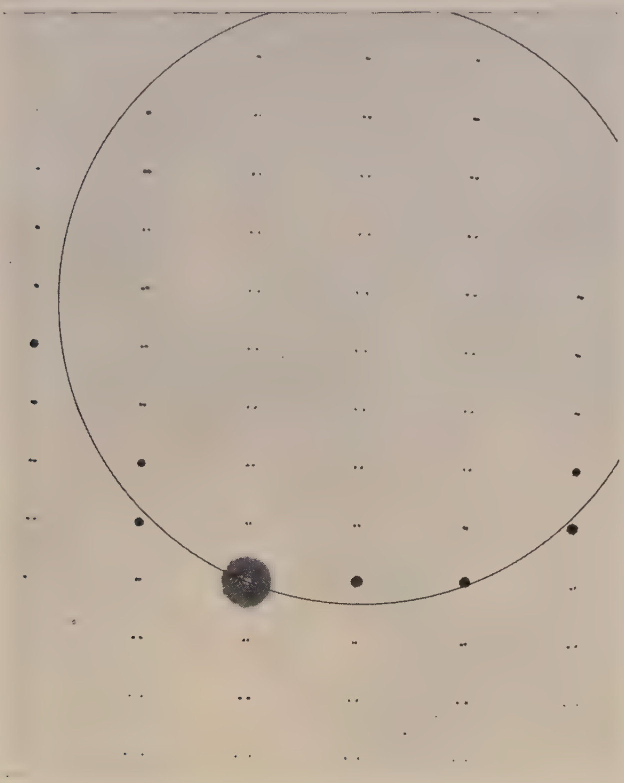


Figure 1. Cross-grating pattern from zinc oxide smoke particle showing doubling of spots. The Laue circle is indicated. A background of rings and spots has been removed by masking.

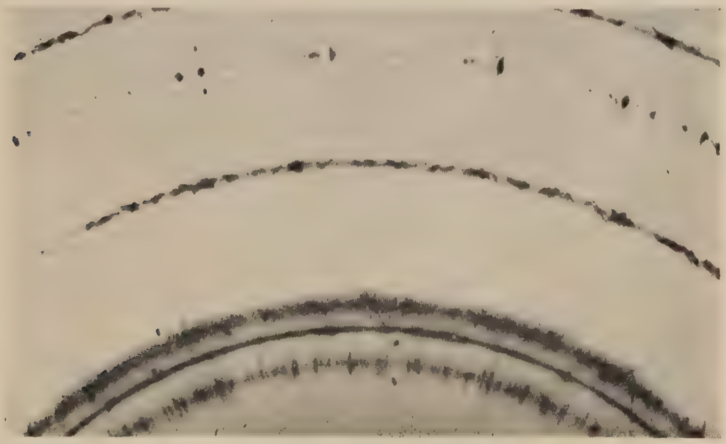


Figure 4. Zinc oxide smoke pattern showing pairs of spots about the $10\bar{1}0$ and $10\bar{1}1$ rings (first and third) and a triplet on the $11\bar{2}0$ ring (fifth).

PLATE I.

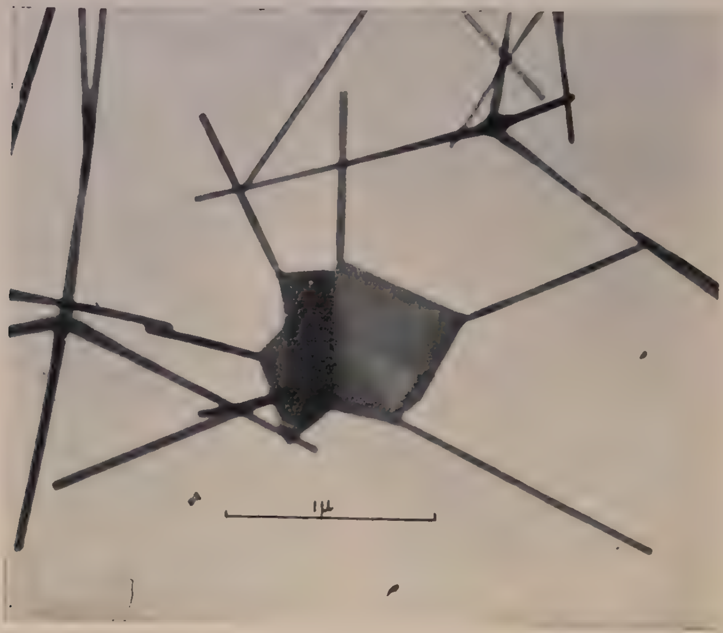


Figure 7. Electron-micrograph of zinc oxide smoke showing a group of spines and sheets.

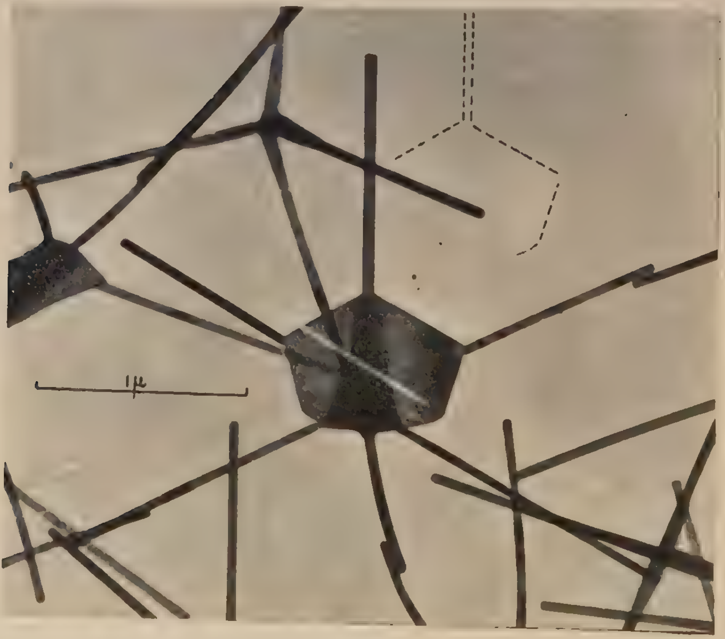
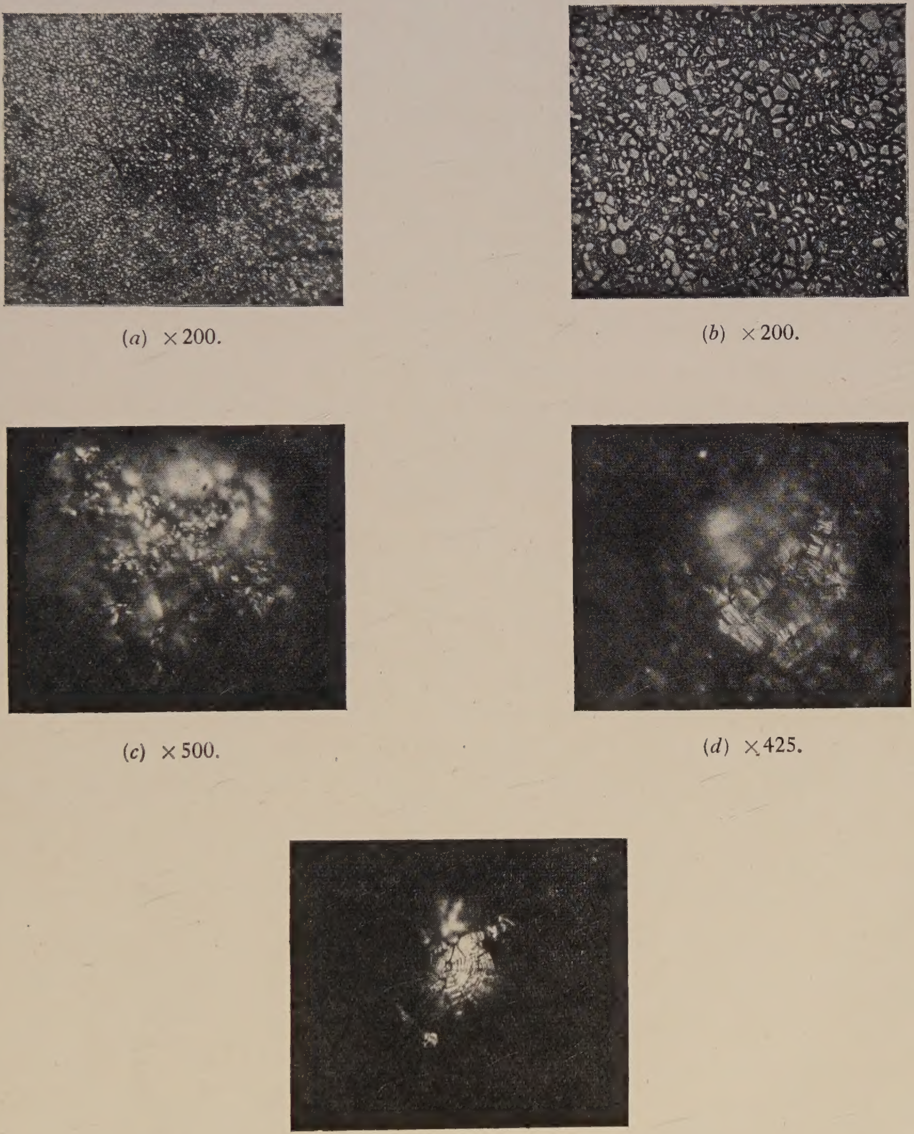


Figure 8. Electron-micrograph of zinc oxide smoke. The outline of a 'dark-field' diffracted image is dotted.

Plate II.



(*a*) ×200.

(*b*) ×200.

(*c*) ×500.

(*d*) ×425.

(*e*) ×425.

Figure 8.

PHYSICAL SOCIETY PUBLICATIONS

Fellows and Student Members of the Society may obtain ONE copy of each publication at the price shown in brackets. In most cases the cost of postage and packing is extra.

Noise and Sound Transmission. Report of the 1948 Summer Symposium of the Acoustics Group of the Physical Society. Pp. 200. In paper covers. 17*s.* 6d. (10*s.* 6*d.*) Postage 6*d.*

Resonant Absorbers and Reverberation. Report of the 1947 Summer Symposium of the Acoustics Group of the Physical Society. Pp. 57. In paper covers. 7*s.* 6*d.* (5*s.*) Postage 6*d.*

The Emission Spectra of the Night Sky and Aurorae, 1948. Papers read at an International Conference held under the auspices of the Gassiot Committee in London in July 1947. Pp. 140. In paper covers. 20*s.* (12*s.* 6*d.*) Postage 6*d.*

The Strength of Solids, 1948. Report of Conference held at Bristol in July 1947. Pp. 162. In paper covers. 25*s.* (15*s.* 6*d.*) Postage 8*d.*

Report of International Conference on Fundamental Particles (Vol. I) *and Low Temperatures* (Vol. II), 1947. Conference held at Cambridge in July 1946. Pp. 200 (Vol. I), pp. 184 (Vol. II). In paper covers. 15*s.* each vol. (7*s.* 6*d.*) Postage 8*d.*

Meteorological Factors in Radio-Wave Propagation, 1947. Report of Conference held jointly with the Royal Meteorological Society in April 1946. Pp. 325. In paper covers. 24*s.* (12*s.*+postage 1*s.*)

Handbook of the 35th Exhibition of Scientific Instruments and Apparatus, 1951. Pp. xi+244. In paper covers. 5*s.* (2*s.* 6*d.*) Postage 1*s.*

Handbook of the 34th Exhibition of Scientific Instruments and Apparatus, 1950. Pp. xii+266. In paper covers. 5*s.* (2*s.* 6*d.*) Postage 1*s.*

Handbook of the 33rd Exhibition of Scientific Instruments and Apparatus, 1949. Pp. 272. In paper covers. 5*s.* (2*s.* 6*d.*) Postage 1*s.*

Catalogue of the 32nd Exhibition of Scientific Instruments and Apparatus, 1948. Pp. 288. In paper covers. 5*s.* (2*s.* 6*d.*) Postage 1*s.* (Half price from 5th April 1949.)

Report on Colour Terminology, by a Committee of the Colour Group. Pp. 56. In paper covers. 7*s.* (3*s.* 6*d.*)

Report on Defective Colour Vision in Industry, by a Committee of the Colour Group. 1946. Pp. 52. In paper covers. 3*s.* 6*d.* (1*s.* 9*d.*+postage 4*d.*)

Report on the Teaching of Geometrical Optics, 1934. Pp. 86. In paper covers. 6*s.* 3*d.* Postage 6*d.*

Report on Band Spectra of Diatomic Molecules, 1932. By W. Jevons, D.Sc., Ph.D. Pp. 308. In paper covers, 25*s.*; bound in cloth, 30*s.* (15*s.*) Postage 1*s.*

Discussion on Vision, 1932. Pp. 327. In paper covers. 6*s.* 6*d.* (3*s.* 3*d.*) Postage 1*s.*

Discussion on Audition, 1931. Pp. 151. In paper covers. 4*s.* (2*s.*) Postage 1*s.*

Discussion on Photo-electric Cells and their Application, 1930. Pp. 236. In paper covers. 6*s.* 6*d.* (3*s.* 3*d.*) Postage 8*d.*

The Decimal Bibliographic Classification (Optics, Light and Cognate Subjects), 1926. By A. F. C. Pollard, D.Sc. Pp. 109. Bound in cloth. 4*s.* (2*s.*) Postage 8*d.*

Motor Headlights, 1922. Pp. 39. In paper covers. 1*s.* 6*d.* (9*d.*) Postage 4*d.*

Report on Series in Line Spectra, 1922. By A. Fowler, C.B.E., Sc.D., F.R.S. Pp. 182. In paper covers. 30*s.* (15*s.*) Postage 8*d.*

A Discussion on the Making of Reflecting Surfaces, 1920. Pp. 44. In paper covers. 2*s.* 6*d.* (1*s.* 3*d.*) Postage 4*d.*

Reports on Progress in Physics. Vol. XIV (1951). Pp. 412. Bound in cloth. 50*s.* (27*s.* 6*d.*) Postage 1*s.*

Reports on Progress in Physics. Vol. XIII (1950). Pp. 424. Bound in cloth. 50*s.* (25*s.*) Postage 1*s.*

Reports on Progress in Physics. Vol. XII (1948–49). Pp. 382. Bound in cloth. 42*s.* (25*s.*) Postage 1*s.*

Reports on Progress in Physics. Vol. XI (1946–48). Pp. 461. Bound in cloth. 42*s.* (25*s.*) Postage 1*s.*

Reports on Progress in Physics. Vols. IV (1937, reprinted 1946) and X (1944–45). Bound in cloth. 30*s.* each. (15*s.*) Postage 1*s.*

The Proceedings of the Physical Society. From Vol. I (1874–75), excepting a few parts which are out of print. Prices on application to Messrs. Wm. Dawson Ltd., 102 Wigmore St., London W.1.

The Transactions of the Optical Society. Vols. 1 (1899–1900) –33 (1931–32), excepting a few parts which are out of print. Prices on application to Messrs. Wm. Dawson Ltd., 102 Wigmore St., London W.1.

Orders, accompanied by remittances, should be sent to

THE PHYSICAL SOCIETY
1 Lowther Gardens, Prince Consort Road, London S.W.7

THE PHYSICAL SOCIETY

MEMBERSHIP

Membership of the Society is open to all who are interested in Physics :

FELLOWSHIP. A candidate for election to Fellowship must as a rule be recommended by three Fellows, to two of whom he is known personally. Fellows may attend all meetings of the Society, are entitled to receive Publications 1 (either Section A or Section B), 4 and 5 below, and may obtain the other publications at much reduced rates.

STUDENT MEMBERSHIP. A candidate for election to Student Membership must be between 18 and 26 years of age and must be recommended from personal knowledge by a Fellow. Student Members may attend all meetings of the Society, are entitled to receive Publications 1 (either Section A or Section B) and 4, and may obtain the other publications at much reduced rates.

Books and periodicals may be read in the Society's Library, and a limited number of books may be borrowed by Fellows and Student Members on application to the Honorary Librarian.

Fellows and Student Members may become members of the *Colour Group*, the *Optical Group*, the *Low Temperature Group* and the *Acoustics Group* (specialist Groups formed in the Society) without payment of additional annual subscriptions.

PUBLICATIONS

1. *The Proceedings of the Physical Society*, published monthly in two Sections, contains original papers, lectures by specialists, reports of discussions and of demonstrations, and book reviews. Section A contains papers mainly on atomic and sub-atomic subjects; Section B contains papers on macroscopic physics.

2. *Reports on Progress in Physics*, published annually, is a comprehensive review by qualified physicists.

3. *The Handbook of the Physical Society's Annual Exhibition of Scientific Instruments and Apparatus.* This Exhibition is recognized as the most important function of its kind, and the Handbook is a valuable book of reference.

4. *The Bulletin*, issued at frequent intervals during the session, informs members of the programmes of future meetings and of the business of the Society generally.

5. *Physics Abstracts* (*Science Abstracts A*), published monthly in association with the Institution of Electrical Engineers, covers the whole field of contemporary physical research.

6. *Electrical Engineering Abstracts* (*Science Abstracts B*), published monthly in association with the Institution of Electrical Engineers, covers the whole field of contemporary research in electrical engineering.

7. *Special Publications*, critical monographs and reports on special subjects prepared by experts or committees, are issued from time to time.

MEETINGS

At approximately monthly intervals throughout each annual session, meetings are held for the reading and discussion of papers, for lectures, and for experimental demonstrations. Special lectures include: the *Guthrie Lecture*, in memory of the founder of the Society, given annually by a physicist of international reputation; the *Thomas Young Oration*, given biennially on an optical subject; the *Charles Chree Address*, given biennially on Geomagnetism, Atmospheric Electricity, or a cognate subject; and the biennial *Rutherford Memorial Lecture.* Meetings are generally held each year at provincial centres, and from time to time meetings are arranged jointly with other Societies for the discussion of subjects of common interest.

Each of the four specialist Groups holds about five meetings in each session.

SUBSCRIPTIONS

Fellows pay an Entrance Fee of £1 1*s.* and an Annual Subscription of £3 3*s.* ; Student Members pay only an Annual Subscription of 15*s.* Second Section of *Proceedings* 30*s.* No entrance fee is payable by a Student Member on transfer to Fellowship.

Further information may be obtained from the Secretary-Editor
at the Offices of the Society :

1 LOWTHER GARDENS, PRINCE CONSORT ROAD, LONDON S.W.7
Telephone : KENsington 0048, 0049

Printed by TAYLOR AND FRANCIS, LTD., Red Lion Court, Fleet Street, London E.C.4.